ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

160
CONTRIBUTIONS

BAYESIAN METHODS APPLIED FOR ECOSYSTEM MODEL CALIBRATION AND UNCERTAINTY SOURCE ESTIMATION

JARMO MÄKELÄ

FINNISH METEOROLOGICAL INSTITUTE
CONTRIBUTIONS
No. 160

Bayesian methods applied for ecosystem model calibration and uncertainty source estimation

Jarmo Mäkelä

Doctoral Programme in Mathematics and Statistics
Department of Mathematics and Statistics
Faculty of Science
University of Helsinki

Academic dissertation presented for
the degree of Doctor of Philosophy

To be presented for public examination with the permission of
the Faculty of Science of the University of Helsinki
in the Auditorium Brainstorm at the Finnish Meteorological Institute, Helsinki,
on 7th of February 2020 at noon.

Finnish Meteorological Institute
Helsinki, 2020

Author's contact: Climate System Research, Finnish Meteorological Institute
PL503, FI-00101 Helsinki, Finland
jarmo.makela@fmi.fi
https://orcid.org/0000-0002-8788-3939

Supervisors: Docent Tuula Aalto, Ph.D.
Department of Physics, University of Helsinki
Climate System Research, Finnish Meteorological Institute

Professor Jukka Corander, Ph.D.
Department of Mathematics and Statistics,
University of Helsinki

Tiina Markkanen, Ph.D.
Climate System Research, Finnish Meteorological Institute

Reviewers: Professor Kari Auranen, Ph.D.
Department of Mathematics and Statistics,
University of Turku

Associate Professor Ari Laurén, Ph.D.
School of Forest Sciences,
University of Eastern Finland

Opponent: Professor Dr. Florian Hartig
Faculty for Biology and Preclinical Medicine,
University of Regensburg

Custos: Associate Professor Matti Pirinen, Ph.D.
Department of Mathematics and Statistics,
University of Helsinki

ISBN 978-952-336-094-5 (paperback)
ISSN 0782-6117
Edita Prima Oy
Helsinki 2020

ISBN 978-952-336-095-2 (PDF)
https://doi.org/10.35614/isbn.9789523360952
Helsinki 2020

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Published by	Series title, number and report code of publication
Finnish Meteorological Institute	FMI Contributions, 160, FMI-CONT-160
(Erik Palménin aukio 1), P.O. Box 503	
FIN-00101 Helsinki, Finland	January 2020

Author	ORCID iD
Jarmo Mäkelä	0000-0002-8788-3939

Title

Bayesian methods applied for ecosystem model calibration and uncertainty source estimation

Abstract

How significant are different uncertainty sources when simulating the future state of the ecosystem in Finland?

In this thesis, we examine this question and provide some answers to this broad topic by simulating 21st century ecosystem conditions with a land-ecosystem model called JSBACH. The results are also compared to similar simulations performed by another model called PREBAS. We consider four different sources of uncertainty that are related to 1) the model that is used to generate the future conditions; 2) future climate used to drive the model, represented by an ensemble of CMIP5 simulations; 3) RCP scenarios that depict the rising atmospheric CO_2 concentration and; 4) forest management actions.

Before running the simulations described above, we calibrated and validated the JSBACH model extensively on different temporal resolutions and with multiple model modifications. These hindcasting calibrations were performed with two Bayesian approaches: the adaptive Metropolis algorithm and the adaptive population importance sampler. The calibrations resulted in a sufficient model setup and satisfactory parameter distributions. These were used to represent the JSBACH model uncertainty in the 21st century simulations.

Canonical correlation analysis was used to gleam the impact of the different uncertainty sources on multiple groups of ecosystem variables. The results are summarised via the use of redundancy indices that yield varied impacts. Overall, forest management actions and RCP scenarios tend to dominate the uncertainties towards the end of the century, but the effect of climate models and parameters should not be overlooked especially since a more detailed examination revealed that their impact was not fully captured.

Publishing unit	Classification (UDC)
Climate System Research	

Keywords

Bayesian statistics, boreal, calibration, climate change, data-assimilation, ecosystem, estimation, hindcasting, importance sampling, Markov chain Monte Carlo, MCMC, Metropolis-Hastings, uncertainty

ISSN and series title	ISBN
0782-6117	978-952-336-094-5 (paperback)
Finnish Meteorological Institute Contributions	978-952-336-095-2 (pdf)

DOI	Language	Pages
10.35614/isbn.9789523360952	English	116

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Julkaisija	Julkaisun sarja, numero ja raporttikoodi
Ilmatieteen laitos (Erik Palménin aukio 1) PL 503, 00101 Helsinki	Contributions, 160, FMI-CONT-160 Tammikuu 2020

Tekijä	ORCID iD
Jarmo Mäkelä	0000-0002-8788-3939

Nimeke

Bayesian methods applied for ecosystem model calibration and uncertainty source estimation

Tiivistelmä

Kuinka merkittäviä ovat eri epävarmuuslähteet arvioitaessa metsäekosysteemien tulevaisuutta Suomessa?

Tässä väitöskirjassa tarkastellaan edeltävää kysymystä mallittamalla metsäekosysteemien tilaa 2100-luvulle. Mallitukseen käytetään maa-ekosysteemimalli JSBACH:ia ja arvioita verrataan vastaaviin PREBAS-mallin tuloksiin. Tarkasteltavat epävarmuuslähteet voidaan jakaa 1) mallien sisäiseen epävarmuuteen; 2) mallien ajamiseen käytettäviin ilmastopakotteisiin, jotka pohjaavat CMIP5 simulaatioihin; 3) RCP-päästöskenaarioihin, jotka edustavat ilmakehän hiilidioksidipitoisuuden nousua sekä; 4) valittuun metsänhoitosuunnitelmaan.

Edellä esitettyjen tulevaisuussimulaatioiden toteuttamiseksi JSBACH-malli kalibroitiin ja validoitiin käyttäen 10 paikallista mittausasemaa boreaalisella vyöhykkeellä. Tarkasteluun sisällytettiin eri aikaresoluutioita ja useita mallin rakenteellisia muutoksia, jotta mallin tuottama transpiraatio, evaporaatio ja hiilensidonta vastaisivat paremmin vastaavia havaintoja. Kalibrointiin käytettiin kahta eri Bayesilaista menetelmää: adaptive Metropolis sekä adaptive population importance sampler -algoritmeja. Kalibrointiprosessissa ilmenneitä parametrien välisiä riippuvuuksia, identifioituvuutta ja merkitsevyyttä analysoitiin perusteellisesti. Lisäksi lopullisia parametrijakaumia ja -arvoja verrattiin useisiin kirjallisuuslähteisiin. Validointi toteutettiin sekä kalibroinnista riippumattomilla että erillisten mittausasemien havainnoilla. Näiden tarkastelujen pohjalta muodostettiin tulevaisuussimulaatioihin soveltuva mallirakenne ja parametrijakaumat, jotka kuvaavat JSBACH-mallin sisäistä epävarmuutta.

Tulevaisuussimulaatioiden epävarmuuslähteiden vaikuttavuutta arvioitiin kanonisen korrelaatioanalyysin pohjalta. Metsäekosysteemien tilaa kuvaavia indikaattoreita tutkittiin em. analyysissa sekä kaikkia kerralla että jaoteltuna eri ryhmiin vaikuttavuutensa perusteella. Voimakkaimmaksi epävarmuuslähteeksi vuosisadan lopussa nousivat sekä metsänhoitosuunnitelmat että päästöskenaariot, joita seurasivat järjestyksessä ilmastopakotteet ja mallien sisäinen epävarmuus.

Julkaisijayksikkö	Luokitus (UDK)
Ilmastojärjestelmätutkimus	519.226, 519.233.5, 551.588.6

Asiasanat

Bayes, boreaalinen, data-assimilaatio, ekosysteemi, epävarmuus, ilmastonmuutos, Markovin ketju, MCMC

ISBN	ISSN ja avainnimeke	
978-952-336-094-5 (paperback), 978-952-336-095-2 (pdf)	0782-6117	Contributions

DOI	Kieli	Sivumäärä
10.35614/isbn.9789523360952	Englanti	115

Acknowledgements

The work related to this thesis has been ongoing for several years and there have been many steps, triumphs and setbacks on the way. I have had the opportunity to largely steer the work towards waters that I perceived as interesting, of which I would like to express my gratitude to my supervisor and head of the Carbon Cycle Modelling group at FMI, Doc. Tuula Aalto. She has had the patience to guide me e.g. through several months of wasted computer simulations. Similarly, I would like to express my gratitude to my other supervisor at the FMI, Dr. Tiina Markkanen, who has always been willing to fill in my blanks on ecosystem process and to speculate and discuss the results of my work. I would like to also thank my supervisor at the University, Prof. Jukka Corander for steering me towards importance sampling and for trusting me to get on with my research.

I also want to recognise the efforts of Prof. Kari Auranen and Prof. Ari Laurén for reviewing and improving the thesis manuscript – especially the comments of the former on the mathematical notation and presentation were invaluable. Similarly, I want to thank Prof. Dr. Florian Hartig for accepting the position of Opponent at the public examination and Prof. Matti Pirinen for agreeing to act as the Custos. Additionally, I am grateful to Prof. Sangita Kulathinal for the thankless job of handling the official University bureaucracy and paperwork during this process.

Much of this work would not have been possible, or it would be of much poorer quality, without the efforts of my co-authors. Their comments and insight greatly improved the articles that constitute the basis of this thesis. They have my praise as do my other colleagues at FMI for providing a supportive and stimulating work environment and occasional leisure activities.

I am also grateful for my family and friends that are not associated with my day-to-day research. They have both supported the work but also kept me grounded and provided a much needed wider perspective in all matters. Their efforts have likely kept me relatively tolerable as a human being.

Lastly, I want to acknowledge the support this work has received from the Jenny and Antti Wihuri Foundation.

Helsinki, January 2020
Jarmo Mäkelä

Definition of abbreviations and symbols

JSBACH	Land surface component of the Earth System model of the Max Planck Institute for Meteorology (MPI-ESM)
PREBAS	Forest model system (PRELES, CROBAS, YASSO07)
AM	Adaptive Metropolis (algorithm)
APIS	Adaptive population importance sampler
MCMC	Markov chain Monte Carlo
$\mathcal{M}$	Model or observation operator
$\mathcal{L}$	Likelihood function
Z	Model evidence or partition function
pdf	Probability density function
$\pi(\boldsymbol{\theta})$	Parameter prior pdf
$p(\boldsymbol{\theta}\|\mathbf{x})$	Parameter posterior pdf
$q(\boldsymbol{\theta}\|\mathbf{x})$	Parameter proposal or sampling pdf
$\mathbf{x}$	State vector
$\mathbf{y}$	Observation vector
$\boldsymbol{\theta}$	Parameter vector
$\boldsymbol{\theta}^j$	j-th component of the parameter vector
$[\ \ ]_i$	i-th vector, distribution or draw from the parameter space
CCA	Canonical correlation analysis
KDE	Kernel density estimate
PCA	Principal component analysis
EC	Eddy covariance
ET	Evapotranspiration
GPP	Gross primary production
LAI	Leaf area index
$[\ \ ]_{obs/mod}$	Underscore depicting observed / modelled variable

List of publications with author contribution

This dissertation consists of an introductory part and three research articles, denoted by Roman numerals **I**–**III**:

I Mäkelä, J., Susiluoto, J., Markkanen, T., Aurela, M., Järvinen, H., Mammarella, I., Hagemann, S., and Aalto, T.: Constraining ecosystem model with adaptive Metropolis algorithm using boreal forest site eddy covariance measurements, Nonlin. Processes Geophys., 23, 447-465, doi:10.5194/npg-23-447-2016, 2016.

II Mäkelä, J., Knauer, J., Aurela, M., Black, A., Heimann, M., Kobayashi, H., Lohila, A., Mammarella, I., Margolis, H., Markkanen, T., Susiluoto, J., Thum, T., Viskari, T., Zaehle, S., and Aalto, T.: Parameter calibration and stomatal conductance formulation comparison for boreal forests with adaptive population importance sampler in the land surface model JSBACH, Geosci. Model Dev., 12, 4075–4098, doi:10.5194/gmd-12-4075-2019, 2019.

III Mäkelä, J., Minunno, F., Aalto, T., Mäkelä, A., Markkanen, T., and Peltoniemi, M.: Uncertainty sources in simulated ecosystem indicators of the 21st century climate change, Biogeosciences Discuss., doi:10.5194/bg-2019-395, 2019.

I am the principal author of all three articles included in this thesis. Each article contains a section on author contributions. The first authorship in article **III** is shared between myself and F. Minunno. In article **III**, I am responsible for the JSBACH simulations and F. Minunno for the PREBAS results. During the thesis work, I have e.g. worked on both model and algorithm development and implementation; planned and executed computer simulations and model calibrations; identified suitable measurement sites for the simulations and processed the relevant data; analysed the results and prepared the article manuscripts.

Contents

1 Introduction

Predicting future climate conditions is occasionally frustrating due to many different error sources, such as insufficient or too simplistic models or inaccurate model parameter estimates. These deficiencies can substantially affect the simulated conditions and therefore identifying and resolving them is of interest to the modelling community in order to provide more realistic projections. Future climate conditions are commonly assessed by multimodel ensembles, of which the most well known is the Coupled Model Intercomparison Project (CMIP) that provides knowledge for the Intergovernmental Panel on Climate Change (IPCC).

The ensemble approach in CMIP relies on models compensating each others' weaknesses, but it does not remove the need for individual model improvements and calibration. Mueller and Seneviratne (2014) evaluated the realism of simulated evapotranspiration (ET), precipitation and temperature in the CMIP5 ensemble on continental areas. CMIP5 is the latest, completed phase (5th) of the CMIP project that mostly focused on atmospheric variables. The results revealed systematic ET biases in CMIP5 model simulations with an overestimation in most regions. Boé and Terray (2008) have shown that this has significant implications for climate change projections.

Research detailed in this thesis examines model parameter calibration and relative uncertainty source estimation in JSBACH, an ecosystem model which is the land component of the Earth System model of the Max Planck Institute for Meteorology (MPI-ESM). MPI-ESM was one of the models included in CMIP5 and the model consortium is currently in the process of producing simulations for CMIP6, introduced in Eyring and Taylor (2016). The CMIP6 has expanded the focus and also includes e.g. land use and more emphasis on land surface processes.

Similarly to results shown by Mueller and Seneviratne (2014), JSBACH has an insufficient response to water limitation in Finland – it overestimates evapotranspiration. This was verified by Peltoniemi et al. (2015a) who estimated the total (plant) gross primary production (GPP) related to carbon uptake of Finnish forests. Summer 2006 saw significantly less rain than what is usual. Another model, called PRELES, was able to transfer the reduced rainfall into lower levels of GPP – this was scientifically expected and confirmed by the observations. However this characteristic was not captured by the JSBACH model.

The overall aim of this thesis is to test the suitability of two different Bayesian calibration methods in improving the modelled ET and GPP fluxes, and to provide estimates of relative uncertainty sources in future predictions of ecosystem indicators of climate change. We consider model parameterisation as one of these uncertainty sources. This component has been neglected in many similar studies (Kuppel et al., 2012).

We begin this process by calibrating the ecosystem model JSBACH in paper **I** with the adaptive Metropolis (AM) algorithm, a representative of Markov chain Monte Carlo (MCMC) samplers. The JSBACH model is calibrated on different temporal resolutions and we examine both the optimisation process and the optimised fluxes.

Subsequently, we chose one of the temporal resolutions to be used in further calibration attempts in paper **II**, utilising a different approach with the adaptive population importance sampler (APIS). The calibrations and model setup in paper **II** aimed to resolve previously identified model weaknesses. Moreover, we asserted a multi-site calibration design on several interchangeable submodel components to reduce site-specificity and to better assess model weaknesses.

Suitable parameter distributions from paper **II** were used to perturb the model in paper **III**, where we examined the relative uncertainty sources in simulated ecosystem indicators of the 21st century climate change. We assessed the ecosystem indicator variability with canonical correlation analysis on grouped ecosystem indicators and presented the results via the use of redundancy indices. The analysis highlights the magnitude of forest management uncertainty and perceives that the impact of parameters, and therefore the model internal variability, may be unaccounted and severely underestimated.

2 Model setup and data

In this thesis, the focus is on the process-based ecosystem model JSBACH (Kaminski et al., 2013), capable of simulating the water, energy and carbon balances of the surface and soil (Roeckner et al., 2003). The examined processes and parameters in paper **I** are related to photosynthesis, soil hydrology and phenology. In paper **II**, we extend this set with multiple stomatal conductance formulations. Descriptions of these parameters, processes and the governed equations are given in papers **I** and **II**. In this thesis, the model behaviour is examined on altogether ten coniferous evergreen sites in the boreal zone.

2.1 JSBACH model

The JSBACH model requires various data sources to function, separated into three distinct categories: *location characteristics*, *state variables* and *meteorological measurements*. The simulations covered in this thesis are all restricted to site-level. Thus for each site, the different data sources represent the same geographical location defined as the footprint area of the eddy covariance (EC) measurement tower located at the site. EC is a technique used to measure and calculate turbulent vertical fluxes within the boundary layer of the atmosphere and the land-surface (Baldocchi et al., 1988).

The *site-level* (or location) *characteristics* are fixed parameters defining e.g. the site coordinates, amount of bare soil and the type and amount of vegetation, elevation and orography, soil depth, field capacity, porosity, heat capacity and heat conductivity.

The *state variables* define the status of the model and they depict both model internal processes as well as observable conditions. These variables are typically initialised to some default values or empty initial storage pools. In order for these variables to depict the site properties reasonably well (e.g. correct soil moisture content and leaf area index), a model spin-up is required. Our spin-up is a couple of decades long and is enough to equilibrate the variables of interest. In contrast, this spin-up would not be enough for e.g. soil carbon pools that would require a spin-up of several thousand years to reach equilibrium.

Since our simulations are restricted to site-level, the model utilises *in-situ measurements* of air temperature, atmospheric carbon dioxide (CO_2) concentration, precipitation, pressure, specific humidity, short- and longwave radiation, potential shortwave radiation and wind speed. These half-hourly values are the only external source of information for the model, and they are used by the model to transition from one state to the next.

The state variables include three specific data streams that are used for model calibration, namely leaf area index (LAI), evapotranspiration (ET) and (plant) gross primary production (GPP). Two of these variables, ET and GPP, are observed at site level using EC flux measurements. The EC method can be used to estimate e.g. the exchange rates of water vapour and carbon dioxide, which are respectively used to derive the ET and GPP observations. The ET and GPP estimates are produced to the modelling community by the site primary investigators (PI's).

LAI depicts the total surface area of leaves (needles for conifers) within a given area of ground.

ET is directly measured by the EC method and it accounts for the movement of water from the ground to the atmosphere. The main components of ET are water evaporation (from soil and other surfaces) and plant transpiration.

GPP represents the amount of carbon (chemical energy) fixed by plants during photosynthesis. The ecosystem also loses carbon due to plant internal metabolism (autotrophic respiration) and as a result of e.g. soil microbial processes (heterotrophic respiration). These carbon sources have to be taken into account when estimating GPP. The EC measurements of CO_2 are used to determine the net ecosystem exchange (NEE) of CO_2 (total carbon exchange). This estimate is then partitioned into GPP and (total) respiration (R) of CO_2 (NEE=GPP+R) by utilising flux partitioning procedures (Kolari et al., 2009; Reichstein et al., 2005).

We also define water use efficiency (WUE) at the ecosystem level as the amount of carbon assimilated during photosynthesis, divided by the water used in this process: WUE=GPP/ET.

2.2 Model calibration variables

In the simulations related to this thesis, the JSBACH land surface consists of site-specific grid-cells split into bare soil and vegetative areas. The vegetative area can be further divided into tiles representing the most prevalent vegetation classes, called plant functional types (PFTs) (Reick et al., 2013) – we use only one PFT, coniferous evergreen trees.

We examine site-level forest dynamics by reproducing the amount of carbon assimilated (GPP) during photosynthesis and the amount of water used in this process (ET). This approach requires that the model is able to simulate e.g. forest growth, leaf shedding and soil water absorption. Here we give a general overview of these processes, focusing on the calibration variables: LAI, ET and GPP. The essential interactions related to these variables are depicted in Fig. 1. More explicit explanations with related equations

can be found in papers **I** and **II** or by Hagemann and Stacke (2015); Kaminski et al. (2013); Reick et al. (2013); Raddatz et al. (2007); Roeckner et al. (2003).

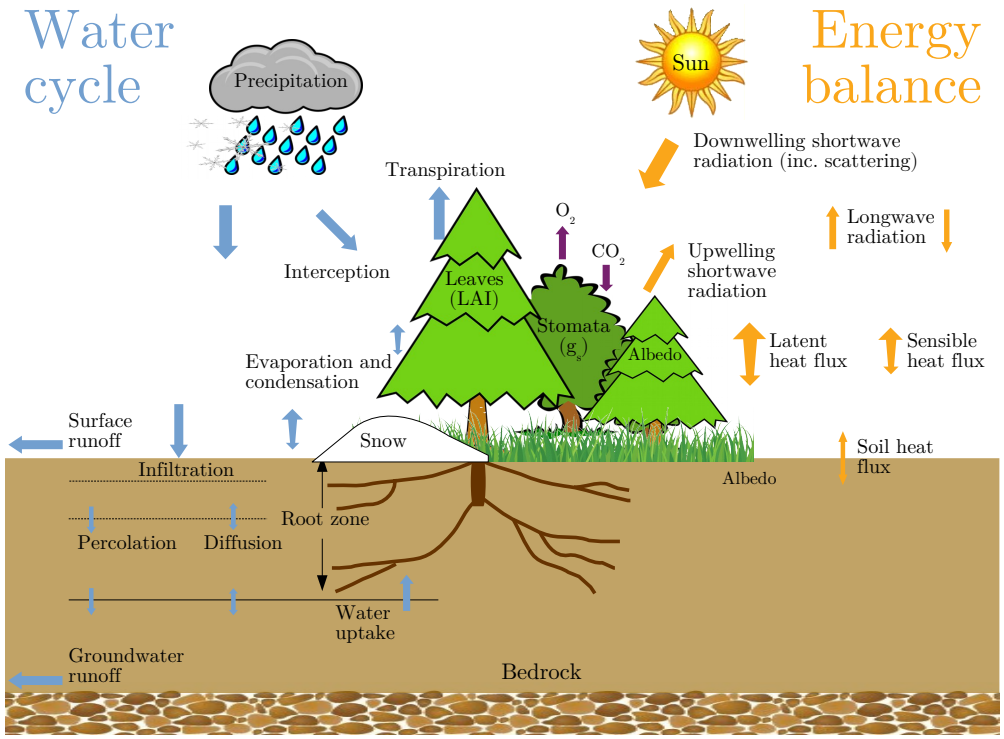

Figure 1: Essential interactions for leaf area index, evapotranspiration and gross primary production. Image adapted from Gao (2016).

2.2.1 Leaf area index

The leaf area index (LAI) is a measure of the plant photosynthetic capability, where plant production (explained in section 2.2.3) is first calculated on leaf-level and then integrated over the whole canopy (LAI). LAI has a clear seasonal cycle in the boreal zone as trees, including conifers, grow more leaves or needles during the spring/summer. The conifers start to gradually shed some of their needles towards the end of the summer. A new needle typically stays on the tree for a couple of years. The seasonal development of LAI follows a logistic function and it is regulated by air temperature and soil moisture with a PFT-specific limiting value for the maximum of LAI. In paper **I**, we calibrated this maximum value at a seasonal level. In paper **II**, we set the maximum to a realistic value and adjusted the site-specific fractions of vegetative area to reproduce the measured site-level maximum of LAI.

2.2.2 Evapotranspiration and soil water

The site-level soil water input in JSBACH is realised via rainfall and snow. Precipitation is divided onto bare soil and the different PFT's, based on their relative area fractions. The rain can be intercepted by the canopy or it can fall on the ground, where it is distributed among abstract "surface" reservoirs. The water in these reservoirs is allowed to evaporate, but any excess is either moved to soil water or removed as runoff.

In paper **I**, the JSBACH soil water is depicted with a "bucket" scheme. This means that the soil water is idealised as a bucket, with a site-specific maximum water content (soil field capacity). The bucket fills up due to rainfall and lose water as drainage, runoff or through plants. Plants can access water within their rooting zone, from where it is transferred to leaves, used for photosynthesis and transpired to the atmosphere. The bucket scheme is also used by the PREBAS model in paper **III**.

Hagemann and Stacke (2015) introduced a multilayer scheme to satisfy deficiencies in the simple bucket model. In papers **II** and **III**, we utilise this scheme applying five layers. This approach allows for a free distribution of layer heights and a more accurate depiction of water movement through the layers.

2.2.3 Plant production and stomatal conductance

Photosynthesis in JSBACH is described by a biochemical photosynthesis model (Farquhar et al., 1980), where radiation (light) is converted into chemical energy (carbohydrates). In essence, this process requires light, water (H_2O) and carbon dioxide (CO_2) to work. Radiation absorption by plants in JSBACH is estimated by a two-stream approximation within a three-layer canopy (Sellers, 1985). Plants extract water from the soil with their roots and CO_2 is captured from the surrounding air.

Availability of soil water induces limitations to photosynthesis that in JSBACH is represented by the soil water stress function (β). This is a linearly increasing function of volumetric soil water content (θ). The stress function is zero below the plant permanent wilting point (θ_{pwp}), and one above the point where transpiration is not limited by soil moisture (θ_{tsp}). Both θ_{pwp} and θ_{tsp} are fractions of the soil field capacity, which is the maximum value for θ.

$$\beta(\theta) = \begin{cases} 1, & \theta \geq \theta_{\mathrm{tsp}} \\ \frac{\theta - \theta_{\mathrm{pwp}}}{\theta_{\mathrm{tsp}} - \theta_{\mathrm{pwp}}}, & \theta_{\mathrm{pwp}} < \theta < \theta_{\mathrm{tsp}} \\ 0, & \theta \leq \theta_{\mathrm{pwp}}. \end{cases} \tag{2.1}$$

The intensity of photosynthesis also depends on multiple external conditions such as temperature, pressure and humidity, all of which affect the plant stomatal behaviour.

The stomata are microscopic holes on the outer layer of the leaf that regulate the plant CO_2 exchange and water loss. The plant production is idealised and calculated on the stomatal level and then integrated over the total leaf area. Various attempts have been made to correctly model stomatal behaviour, and in paper **II** we examine six different stomatal conductance models (the equations are given in Appendix B of paper **II**). The JSBACH default conductance formulation was used in papers **I** and **III**.

> The *Baseline* (Knorr, 1997) and *Bethy* (Knorr, 2000) models are respectively the original and current default conductance formulations in JSBACH. They both first solve the potential conductance, which is then multiplied by the soil water stress function. The Bethy formulation has an additional check for atmospheric moisture limitation.
>
> The *Ball–Berry* version (Ball et al., 1987), and the three variants included in this thesis, all express the stomatal conductance (g_s) according to the following equation:
>
> $$g_s = g_0 + g_1 \frac{\beta A_n}{C_a}. \tag{2.2}$$
>
> The above formulation contains model-specific parameters (g_0, g_1), soil moisture stress function (β), the net assimilation rate (A_n) and the ambient CO_2 concentration (C_a). The main differences arise from what else is included. The classical Ball-Berry version only multiplies g_1 by relative humidity at leaf surface; *Leuning* model (Leuning, 1995) corrects C_a with CO_2 compensation point and vapour pressure deficit; *F&K* formulation (Friend and Kiang, 2005) introduces exponential reduction based on atmospheric humidity; and *USO* model (Medlyn et al., 2011) readjusts the terms and also introduces vapour pressure deficit into the mix.

In paper **II**, the JSBACH model was also modified to use a delayed effect of temperature for photosynthetic activity in spring. This modification was added to restrain the model photosynthesis of conifers – the springtime GPP, in the paper **I** simulations, had been increasing too rapidly at the start of the photosynthetically active season.

2.3 Measurement sites on the boreal zone

The measurement sites used in this thesis are all located on the boreal zone. We surveyed several data repositories for suitable sites to be included in the calibrations described in paper **II**. The main criterion in assessing the sites was that they should represent boreal forests where the dominant species is evergreen needle-leaf trees. We also required that the site EC tower measurements had been operating continuously for several years. The accepted sites, their locations and vegetation zones along with the dominant species information are presented in Fig. 2.

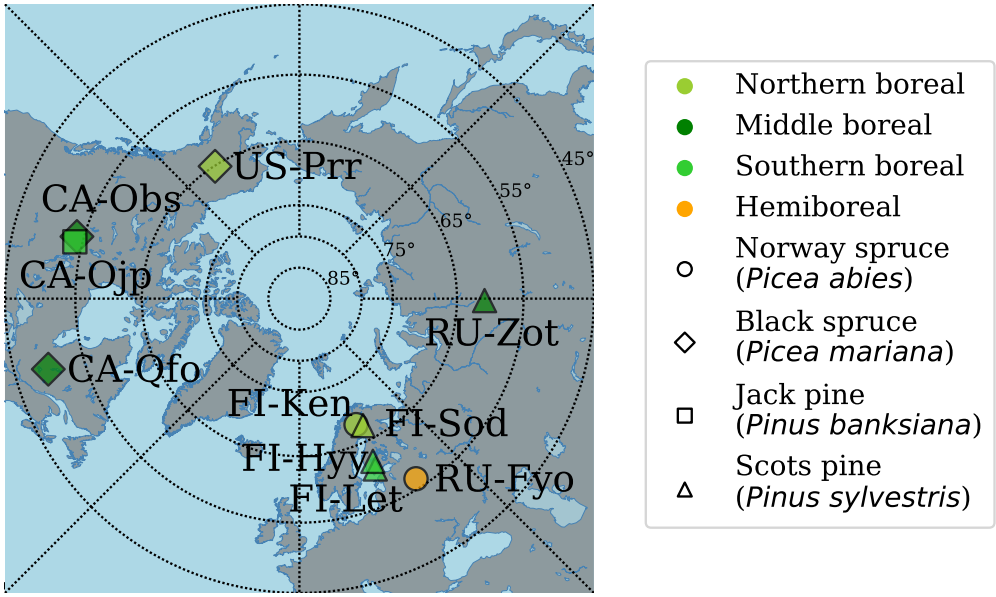

Figure 2: Approximate locations of the measurement sites on a Polar stereographic map with vegetation zone and dominant species information.

All measurement sites have their own distinct features, such as different soil compositions, varying annual precipitation and temperature and seasonal maximum of LAI. These properties are detailed in the papers contained in this thesis and in the reference articles – the properties are taken into account as site-level characteristics in the model. The article references, along with the site names and the name of the person who provided the site-level data, are gathered in Table 1.

Hyytiälä and Sodankylä site measurements are used in paper **I** to calibrate the JSBACH model with different temporal resolutions. All sites in Table 1 are included in paper **II**, where we calibrate and validate the model with multiple stomatal conductance formulations. Hyytiälä and Sodankylä site characteristics are also used in paper **III**, where we investigate uncertainties related to predictive climate simulations in Finland.

2.4 Changing environmental conditions

In papers **I** and **II**, we run hindcasting simulations to calibrate the JSBACH model. This is done to ensure that the model can reproduce the current site-level measurements reasonably well. We cannot guarantee that this approach is enough for the model to correctly simulate the end-of-the-century conditions, but as a process-based model it is suitable for examining robust characteristics of the simulations, e.g. relative responses to different stimuli.

Site ID	Site name and vegetation		Data contact	Reference
CA-Obs	Saskatchewan	MB, BS	Andrew Black	Chen et al. (2006)
CA-Ojp	Saskatchewan	SB, JP	Andrew Black	Chen et al. (2006)
CA-Qfo	Quebec	MB, BS	Hank Margolis	Chen et al. (2006)
FI-Hyy	Hyytiälä	SB, SP	Ivan Mammarella	Kolari et al. (2009)
FI-Ken	Kenttärova	NB, NS	Mika Aurela	Aurela et al. (2015)
FI-Let	Lettosuo	SB, SP	Annalea Lohila	Launiainen et al. (2016)
FI-Sod	Sodankylä	NB, SP	Mika Aurela	Thum et al. (2007)
RU-Fyo	Fyodorkovskoye	HB, NS	Martin Heimann	Launiainen et al. (2016)
RU-Zot	Zotino	MB, SP	Martin Heimann	Kelliher et al. (1998)
US-Prr	Poker Flat	NB, BS	Hideki Kobayashi	Ikawa et al. (2015)

Table 1: Site name and vegetation, data contact and reference paper for specific defitions. (HB - Hemiboreal; MB - Middle boreal; NB - North boreal; SB - South boreal) (BS - Black spruce; JP - Jack Pine; NS - Norway spruce; SP - Scots pine)

The Hyytiälä and Sodankylä site characteristics are used in paper **III** to represent southern and northern Finnish boreal forests, respectively. We keep these characteristics static through the simulations in paper **III**, which means that we do not include such processes as plant competition or changing land cover (Reick et al., 2013). These simulations were made to estimate the contribution of *model parameter uncertainty*, *climate model variability* and *representative concentration pathways (RCP) scenario* to the total uncertainty of certain ecosystem indicators of climate change.

The JSBACH *parameter uncertainty* is represented by a set of 100 parameter vectors, derived from simulations in paper **II**.

Climate model variability is realised via four climate models, belonging to the CMIP5 model ensemble. These climate models reproduce the current climate in Northern Europe well (Lehtonen et al., 2016), but are also reasonably varied over Finland (Ruosteenoja et al., 2016). The climate variables (produced by these climate models) were bias-corrected and down-scaled over Finland. We extracted the required variables from representative grid points and use these data to run the model.

The *RCP*'s refer to an increase in radiative forcing at the end of the century – we examine two pathways, resulting in an increase of 4.5 W/m^2 and 8.5 W/m^2. These pathways are presented in the IPCC AR5 report (IPCC, 2014) and they are used by the different climate model communities to produce the required climate variables (separately for each specific pathway).

Besides JSBACH, similar simulations were made with a forest growth model PREBAS (Valentine and Mäkelä, 2005; Peltoniemi et al., 2015b; Minunno et al., 2019). These

simulations were included because the PREBAS model takes into account forest management, which is one of the most prominent human interactions with forest ecosystems. In paper **III**, two forest management actions were utilised. The business as usual (BAU) scenario follows present forest management recommendations in Finland (Rantala et al., 2011) and the delayed ecosystem logging (DEL) scenario aims for near term carbon sink increase by increasing the minimum harvesting diameter for trees.

3 Bayesian model calibration

We are interested in calibrating model parameters in order to investigate and improve the model performance in various conditions. The modelling setup is defined by the observation equation:

$$\mathbf{y} = \mathcal{M}(\boldsymbol{\theta}, \mathbf{x}) + \mathbf{e}. \tag{3.1}$$

The aim is to reproduce the observations ($\mathbf{y}$) with a model ($\mathcal{M}$). The model produces predictions of the observations with some driving data ($\mathbf{x}$) and the current parameter values ($\boldsymbol{\theta}$). In this thesis, $\mathbf{y}$ represents the calibration variables LAI, ET and GPP; $\mathcal{M}$ is the JSBACH model and $\mathbf{x}$ represents both the model state and the meteorological measurements required to run the model. The parameters of interest are presented in papers **I** and **II** and they generally control idealised model functions that imperfectly emulate the underlying physical processes. Examples of such parameters are: leaf internal $C0_2$ concentration as a fraction of ambient CO_2 (f_{C3}), fraction of volumetric soil water content above which fast drainage occurs (θ_{dr}) and cut-off value in heat sum to determine spring event (T_{alt}).

The residuals ($\mathbf{e}$) depict the model-data mismatch and they form the basis of the parameter likelihood function ($\mathcal{L}$). In this work the full error is treated as Gaussian white noise, so the likelihood function is based on the normal distribution and takes the following form, where c is a constant, σ the volatility of the estimate and N is the number of observations:

$$\mathcal{L}(\mathbf{y}|\boldsymbol{\theta}) = c \prod_{i=1}^{N} \exp\left(-\frac{\mathbf{e}_i^2}{\sigma^2}\right). \tag{3.2}$$

The likelihood function can be used to derive the parameter posterior density function ($p(\boldsymbol{\theta}|\mathbf{y})$). We achieve this by utilising Bayes' rule on conditional probability that enables us to express the posterior pdf of the parameters conditionally on the observed data:

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{\mathcal{L}(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{Z(\mathbf{y})}, \qquad \text{where} \qquad Z(\mathbf{y}) = \int \mathcal{L}(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{3.3}$$

Above, $\pi(\boldsymbol{\theta})$ is the parameter prior pdf and $Z(\mathbf{y})$ is the normalisation constant or model evidence (also known as partition function), calculated as a marginal density of the observations. In practice, it is often impossible to represent the posterior pdf analytically. Hence we turn towards numerical methods and try to approximate the posterior pdf with a finite number of estimates. We will present two different and effective approaches for achieving this: the adaptive Metropolis algorithm used in paper **I** and the adaptive population importance sampler applied in paper **II**.

3.1 Adaptive Metropolis (AM)

The basis of adaptive Metropolis (AM) algorithm (Haario et al., 2001) lies upon discrete-time stochastic processes called Markov chains. These finite chains, when properly generated, asymptotically represent a unique target distribution, namely the parameter posterior pdf in Eq. (3.3). To achieve this, a Markov chain is required to satisfy two properties: *aperiodicity* and *irreducibility*. In addition, the chain is usually constructed to be *reversible* – a usual requirement in (non-adaptive) Monte Carlo simulations.

> *Irreducibility* means that any possible state of the chain can be reached from any other state in a finite number of transitions. Topologically this signifies that the state space is connected. *Aperiodicity* is a requirement that the chain does not repeat itself at fixed intervals – the times of visit to the states are not deterministic. For *reversibility*, it is enough that any pair of states satisfies the *detailed balance* condition. Let $P(\boldsymbol{a} \to \boldsymbol{b})$ denote the conditional probability of the process to make a transition from state $\boldsymbol{a}$ to state $\boldsymbol{b}$, given that it is currently in state $\boldsymbol{a}$ (and likewise define the probability $P(\boldsymbol{b} \to \boldsymbol{a})$ for the reverse transition). Detailed balance means that there exists a probability distribution $(\pi^\star)$ so that the following equation holds for any pair of two states $\boldsymbol{a}$ and $\boldsymbol{b}$:
>
> $$\pi^\star(\boldsymbol{a})P(\boldsymbol{a} \to \boldsymbol{b}) = \pi^\star(\boldsymbol{b})P(\boldsymbol{b} \to \boldsymbol{a}). \tag{3.4}$$
>
> The chain is now constructed so that $\pi^\star$ is the parameter posterior distribution in Eq. (3.3). Thus, we are left with defining such transition probabilities as to satisfy Eq. (3.4).

AM is based on the classical Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970) that satisfies the conditions above. The chains in both of these algorithms consists of a sequence of random draws from the state space (alternatively we call this the parameter space). Assume that the chain is at a state $\boldsymbol{\theta}_t$ and a new state $\boldsymbol{\theta'}$ is randomly generated from a proposal density $q(\boldsymbol{\theta})$ within the state space. The algorithm then either repeats the current state or moves to the new state with (acceptance) probability:

$$a(\boldsymbol{\theta'}, \boldsymbol{\theta}_t) = \min\left(1, \frac{\mathcal{L}(\mathbf{y}|\boldsymbol{\theta'})q(\boldsymbol{\theta}_t|\boldsymbol{\theta'})}{\mathcal{L}(\mathbf{y}|\boldsymbol{\theta}_t)q(\boldsymbol{\theta'}|\boldsymbol{\theta}_t)}\right) \tag{3.5}$$

$$\Rightarrow \min\left(1, \frac{\mathcal{L}(\mathbf{y}|\boldsymbol{\theta'})}{\mathcal{L}(\mathbf{y}|\boldsymbol{\theta}_t)}\right). \tag{3.6}$$

Above, the simplified (Metropolis) expression of Eq. (3.6) holds for the AM algorithm as it utilises symmetric proposal (or sampling) densities $(q(\boldsymbol{\theta}))$. However, the AM incorporates adaptation of the proposal density for the new draws. This is significant

as the adaptation violates the detailed balance of Eq. (3.4). It also implies that AM is not strictly Markovian, since the generation of new draws becomes dependent on all previously drawn samples. Fortunately, the AM has been shown to have the correct ergodic properties (Haario et al., 2001), so it will asymptotically reach the correct target distribution.

The acceptance probabilities in Eq. (3.6) steer the chain towards areas of high probability. The chain represents the target pdf but it can also be used to e.g. maximum a posteriori probability (MAP) estimation or to approximate the parameter expected values:

$$E[\boldsymbol{\theta}|\mathbf{y}] = \int \boldsymbol{\theta} p(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta}. \tag{3.7}$$

The quantities in Eq. (3.7) can be directly assessed with Monte Carlo integration – the parameter expected value will merely be the average value over the chain elements. Additionally, the value of the normalising constant $Z(\mathbf{y})$, presented in Eq. 3.3, is irrelevant during this process.

3.2 Adaptive population importance sampler (APIS)

The adaptive population importance sampler (APIS), introduced by Martino et al. (2015), takes a slightly different approach in gauging the posterior pdf in Eq. (3.3). Similarly to the AM, the APIS algorithm utilises proposal distributions to generate new samples. The fundamental idea in importance sampling is that the samples are weighted based on the target distribution – so certain values of the state space have more impact on the outcome, e.g. estimation of the parameter expected values in Eq. (3.7). This premise is reflected in APIS, which utilises a population of importance samplers (IS) simultaneously. The distribution of the IS expected values is predicted to be around all the modes of the target so in this sense APIS is an estimator of the most "important" regions of the target.

A basic importance sampler (IS) functions by introducing a biased sampling distribution ($q(\boldsymbol{\theta})$) that encourages sampling of "important" values. Multiple samples are drawn from this distribution and then reweighed to form an unbiased estimator of e.g. the parameter expected values in Eq. (3.7) – this equation can be rewritten with the reweighing factor ($r(\boldsymbol{\theta})$) as:

$$E[\boldsymbol{\theta}|\mathbf{y}] = \int \boldsymbol{\theta} r(\boldsymbol{\theta}) q(\boldsymbol{\theta}) d\boldsymbol{\theta}, \qquad \text{where} \qquad r(\boldsymbol{\theta}) = \frac{p(\boldsymbol{\theta}|\mathbf{y})}{q(\boldsymbol{\theta})}. \tag{3.8}$$

In practice, a single IS estimator draws N samples $\boldsymbol{\theta}_i, i \in \{1, ..., N\}$, from a single proposal distribution $q(\mathbf{x})$. The estimator then calculates importance

weights (w_i) for each sample, which enables us to approximate the integral Eq. (3.8) as:

$$E[\boldsymbol{\theta}|\mathbf{y}] \approx \frac{\sum_i w_i \boldsymbol{\theta}_i}{\sum_i w_i}, \qquad \text{where} \qquad w_i = \frac{p(\boldsymbol{\theta}_i|\mathbf{y})}{q(\boldsymbol{\theta}_i)}. \tag{3.9}$$

The simple IS estimator alone is rarely sufficient if the target is even slightly complicated. One classical way of tackling this problem is to join multiple IS estimators together. The simplest approach is to calculate the weights for each of these estimators separately and to normalise the result by the combined sum of all weights. However, this leaves the estimators susceptible to "bad" proposals.

APIS employs multiple IS samplers to assess the target pdf ($p(\boldsymbol{\theta}|\mathbf{y})$) and the normalising constant ($Z(\mathbf{y})$), but suppresses the bad proposals by utilising a deterministic mixture approach (Veach and Guibas, 1995; Owen and Yi, 2000). Each sample θ_{ij} is evaluated with respect to all sampling distributions ($q_j, j \in \{1, ..., M\}$), and weighed by the number of samples drawn from that proposal (N_j). This is equivalent to joining the normalised proposal densities together and evaluating the joint pdf. The weights calculated by the deterministic mixture are presented as:

$$w_{ij} = \frac{p(\boldsymbol{\theta}_{ij}|\mathbf{y})}{\sum_j \left(\frac{N_j}{\sum_k N_k}\right) q_j(\boldsymbol{\theta}_{ij})}. \tag{3.10}$$

The parameter expectation values and the normalising constant in Eq. (3.8) can now be estimated by Monte Carlo integration using weights calculated by Eq. (3.10). These estimates can be adjusted after each individual draw without any need to recalculate or adjust past estimates.

In addition, APIS changes the location parameters of the proposal distributions periodically in order to better sample the more "important" regions of the parameter space. We also simultaneously adapt the shape parameters of the proposal distribution, which is not normally part of the APIS procedure. We utilise the self-normalising AMIS estimators by Cornuet et al. (2012) to adapt the diagonal elements of the proposal covariance matrix. The location ($\boldsymbol{\mu}_j$) and shape ($\mathbf{C}_j$) parameters of each proposal (q_j) are updated using only samples drawn from the same proposal distribution:

$$\boldsymbol{\mu}_j = \frac{\sum_{i=1}^{N_j} w_{ij} \boldsymbol{\theta}_{ij}}{\sum_{i=1}^{N_j} w_{ij}}, \qquad \mathbf{C}_j = \frac{\sum_{i=1}^{N_j} w_{ij} (\boldsymbol{\theta}_{ij} - \boldsymbol{\mu}_j)(\boldsymbol{\theta}_{ij} - \boldsymbol{\mu}_j)^T}{\sum_{i=1}^{N_j} w_{ij}}. \tag{3.11}$$

The final configuration of the APIS proposal means ($\boldsymbol{\mu}_j, j \in \{1, ..., M\}$) is expected to localise well the modes of the target, the parameter posterior distribution $p(\boldsymbol{\theta}|\mathbf{y})$.

3.3 Application of AM and APIS

The AM algorithm operates by repeatedly drawing single samples from the parameter space. This process can be time consuming if the parameter space is high-dimensional or the evaluation of the samples is costly because the likelihood generation (model run) takes a long time. The APIS approach significantly reduces the need for sequential draws as most samples can be drawn simultaneously. We will next present a descriptive comparison on how these algorithms were applied in papers **I** and **II**.

1. The initialisation steps of the AM and the APIS algorithms are very similar. In paper **I**, AM is set up with 8 independent parallel chains, and in paper **II**, APIS utilises 40 simultaneous importance samplers. Each chain or sampler has a random starting location in the parameter space and for APIS we use truncated Gaussian distributions for proposal pdf's.

2. In one iteration of AM, each chain runs the JSBACH model once with the currently proposed parameter configuration. The algorithm then calculates the parameter likelihood and accepts or rejects the draw according to Eq. (3.6). Likewise, during one iteration of APIS we draw 50 samples with each IS sampler independently. These draws are then evaluated and reweighted as presented in Eq. (3.8).

3. AM uses all chains to adjust the sampling distribution, shared by all chains. The adjusting starts after a reasonable number of draws from step 2 has been accepted and continues with increasing intervals. The first half of all chains are always discarded as "warm-up" to reduce any initial-state bias.

4. APIS utilises the 50 reweighted draws from step 2 (for each IS sampler separately) to calculate new location and shape parameters. The locations are automatically accepted as there are no rejection criteria. Additionally, all draws in APIS are used to calculate "global" estimates of the parameter expected values. This process utilises the deterministic mixture approach Veach and Guibas (1995); Owen and Yi (2000) and it is fully iterative – there is no need for any recalculations as the previous estimates are directly adjusted (no information is lost either).

5. Both AM and APIS algorithms are returned to step 2, unless we have reached the end of the sampling process.

AM chains are iteratively updated and the sampling distribution is occasionally adjusted. The sample size in APIS is larger (it is not a Markov chain method) and the focus in this work is on the evolution of the locations of the sampling distributions. We utilise APIS with 40 simultaneous proposal distributions and their location parameters are expected to localise well the modes of the target. The deterministic mixture ensures the stability of the estimation of the parameter expected values.

Due to a coding error, some of the original APIS calibration results were not usable (4 out of 6 stomatal conductance models). These calibrations yielded approximately correctly shaped target distributions that were ultimately rerun with a limited number of parameter draws. Thus, after each APIS calibration we utilised a custom stochastic optimiser to sample the parameter space and locate the optimal set of parameter values.

3.4 Likelihood or cost functions

The Bayesian framework requires a likelihood function, defined in Eq. (3.2), which optimally combines model and observational errors. Unfortunately, the JSBACH model error is unknown. The random observational error associated with the EC measurements is often assumed Gaussian but can be more accurately approximated by a symmetric exponential distribution (Richardson et al., 2006). It increases linearly with the magnitude of the flux, with a standard deviation typically less than 20 % of the flux (Rannik et al., 2016; Richardson et al., 2008).

The model is calibrated with ET and GPP measurements, and in one instance also with seasonal maximum of LAI (as presented below). As explained in section 2.1, the flux partitioning used in generating the observed GPP tends to remove most of the CO_2 flux instabilities. Similar instabilities are still present in the ET observations and we have only a rough idea about the general flux error statistics. Furthermore, plausible pointwise errors would also require minimum precision that are instrument- and site-specific (although we could e.g. interpret minimum precision as the variance in stable wintertime measurements). Due to these issues and for the sake of simplicity, we opted to use the (site-specific) observational mean as the point-wise error term in Eq. (3.2).

Because of this more straightforward approach, we call our objective functions cost functions in papers **I** and **II**. They are formulated as the negative log-likelihoods of Eq. (3.2) and presented collectively below. Information regarding the usage of these functions is given in Table 2 at the start of section 4.

$$cf_1 = \left(\frac{\overline{\mathrm{ET}}_{mod} - \overline{\mathrm{ET}}_{obs}}{\overline{\mathrm{ET}}_{obs}}\right)^2 + \left(\frac{\overline{\mathrm{GPP}}_{mod} - \overline{\mathrm{GPP}}_{obs}}{\overline{\mathrm{GPP}}_{obs}}\right)^2 + \left(\frac{\max(\mathrm{LAI}_{mod}) - \max(\mathrm{LAI}_{obs})}{\max(\mathrm{LAI}_{obs})}\right)^2 \tag{3.12}$$

$$cf_2 = \overbrace{\frac{1}{N_{\mathrm{ET}}} \sum \left(\frac{\mathrm{ET}_{mod} - \mathrm{ET}_{obs}}{\overline{\mathrm{ET}_{obs}}}\right)^2}^{\mathrm{NMSE_{ET}}} + \overbrace{\frac{1}{N_{\mathrm{GPP}}} \sum \left(\frac{\mathrm{GPP}_{mod} - \mathrm{GPP}_{obs}}{\overline{\mathrm{GPP}_{obs}}}\right)^2}^{\mathrm{NMSE_{GPP}}} \tag{3.13}$$

$$cf_3 = (1 - r^2_{\mathrm{ET}})\mathrm{NMSE}_{\mathrm{ET}} + (1 - r^2_{\mathrm{GPP}})\mathrm{NMSE}_{\mathrm{GPP}} \tag{3.14}$$

In addition to the simplified error formulation, we normalise the variable residuals with the amount of observations in cost functions (3.13) and (3.14). The normalisation was originally used in paper **I** to ensure numerical stability. It was left unchanged in paper **II** to enable better comparison with results by Knauer et al. (2015), who used the same likelihood formulation and from whom we adopted the stomatal conductance formulations. Additionally, the normalisation in paper **II** keeps the resulting estimate unbiased towards any site as the number of observations per site varies.

The normalisation broadens the shape of the target distribution, but in paper **I** this is not an issue as we examine the parameter MAP estimate, identifiability and correlations that are unaffected by this choice. The APIS algorithm in paper **II** functions by continuously improving the parameter expected value estimates. The expected values are also unaffected by the spread of the target and as we also adapt the shape of the proposal distributions, the effect of normalisation (wider target distribution) is considerably mitigated. The normalisation does affect the speed of convergence of the algorithm, but this is merely a hindrance and should not be viewed negatively towards APIS.

4 Model simulations

The majority of this section consists of descriptions of JSBACH model calibrations and ecosystem indicator simulations, but we will also address some of the methods used to analyse the simulations at the end of the section. Information regarding JSBACH model calibrations has been gathered in Table 2. The premise of the calibrations was to ensure that the model reproduces site-level measurements of GPP and ET. We chose to approach this problem by including LAI in the calibration as this quantity represents the photosynthetic capability of the ecosystem and controls the total amount of transpiration.

#	Paper	Sampler	(chains×samples)	No. parameters	Calibration	Sites
1	**I**	AM	8×20000	12	cf_1: seasonal	1+1
2	**I**	AM	8×20000	12	cf_2: daily	1+1
3	**I**	AM	8×20000	12	cf_2: half-hourly	1+1
4	**II**	APIS	40×5000	18–21	cf_2: daily	6+4
5	**II**	opt.	8×500	5–6	cf_3: drought	1+0

Table 2: Information on five different calibrations: which papers they were used in, what was the composition of the used algorithm or optimiser, how many parameters were calibrated, which cost function was in effect and how many sites were used for calibration and validation.

The JSBACH model default parameterisation reflects an intent to run global or regional simulations. The default values may not correctly depict site-level conditions, as is the case with the maximum (all-sided) LAI for evergreen needle-leaf trees (the default value is 5 m^2/m^2). This value is used as a limiting factor in the logistic equation for the growth of LAI. In ideal conditions, the modelled canopy LAI can reach approximately 80 % of the maximum value. This is considerably less than measured maximum of LAI at certain sites, e.g. Hyytiälä has an observed maximum LAI of 6.5 m^2/m^2 (Kolari et al., 2009).

4.1 Assessment of three temporal resolutions

The AM algorithm was used to calibrate the JSBACH model with three different temporal resolutions that are represented in Table 2 by calibrations 1–3. The model is calibrated and validated with the Hyytiälä measurements, while Sodankylä is used as an independent validation site. The calibrations are detailed in paper **I**, where we initially calibrated the model (calibration 1) at a seasonal level to ensure a proximate match with the observed amounts of ET, GPP and seasonal maximum of LAI. We fixed the LAI-related parameters based on this calibration and re-adjusted the maximum of LAI for Sodankylä.

After the initial tuning, we calibrated the model separately with half-hourly and daily measurements and verified that the eight independent MCMC chains per calibration converged towards the same target distribution. We examined the resulting parameter cross-correlations and checked the absence of non-linear dependencies between the parameters via kernel density estimates (KDE; Rosenblatt, 1956; Parzen, 1962). We also performed principal component analysis (PCA; Pearson, 1901) for the parameter chains of both calibrations and calculated an additional parameter effectiveness measure – these analyses were made to identify the most important parameters in the calibration process.

We examined the model time series and average diurnal cycles of GPP and ET with the model default parameter values and all calibrated parameter sets (seasonal, daily, half-hourly). We also calculated the model bias and root mean squared error (RMSE) estimates for both sites. Additional focus was given to a drought event at the Hyytiälä site – all parameter sets performed equally badly. Overall, we were able to moderately improve the model performance but more importantly we could assess the parameter identifiability and model deficiencies, such as too early rise in springtime GPP and insufficient drought response.

4.2 Comparison of six stomatal conductance formulations

As a result of the temporal calibration attempts, we redesigned the approach in paper **II** and chose daily values as the basis of our subsequent calibration metric. Following Knauer et al. (2015), we utilised altogether six different stomatal conductance formulations to better examine the effect of external conditions on photosynthesis and transpiration. We updated the model to include a multilayer soil moisture scheme (Hagemann and Stacke, 2015) and introduced an experimental limitation to carbon assimilation under soil water stress (Egea et al., 2011). These modifications were made to improve the model's representation of soil water processes and responses to drought. Additionally, we adapted a delayed effect of temperature to photosynthesis (Mäkelä et al., 2004; Kolari et al., 2007) with the purpose of correcting the erroneous increase in springtime GPP, observed in the previous simulations.

Each of the JSBACH calibrations, utilising one of the six stomatal conductance formulations, is of type 4 in Table 2. They are all implemented by the APIS algorithm and with a similar cost function to daily and half-hourly calibrations in section 4.1. Before calibrating the model, we fixed the maximum value for LAI and adjusted the site-specific bare soil and vegetative area fractions to reproduce the measured site-level maximum of LAI. This is in contrast to the approach with different temporal calibrations above, where we adjusted the maximum LAI for Hyytiälä and Sodankylä separately. This new procedure enabled the model to be run and calibrated for all sites simultaneously.

The convergence of the APIS calibration in paper **II** was verified with Gelman-Rubin diagnostic tests (Gelman and Rubin, 1992). We also tested the stability of the APIS global estimates of the parameter expected values. Similarly to the analysis in paper **I**, we visualise the parameter distributions with KDE and calculate parameter effectiveness measures. The JSBACH model performance, with each stomatal conductance formulation, is evaluated by the slope of the regression line (b) and the coefficient of determination (r^2) of both ET and GPP on daily temporal resolution.

The delayed effect of temperature for photosynthesis corrected the timing in the springtime increase in GPP, and the additional limitation to carbon assimilation under soil water stress was only effective during a separate drought event optimisation. This is the same drought event that was examined in paper **I**. The optimisation was performed with a simple stochastic optimiser and a restricted parameter set for all stomatal conductance formulations. We assessed the daily time series with the same metrics (b and r^2) as above, but also inspected a detailed water use efficiency (WUE) response during the drought.

4.3 Ecosystem indicators of the 21st century climate change

The JSBACH model simulations in paper **III**, producing a number of ecosystem indicators, were straightforward to run as we had a prescribed set of different input sources, presented in section 2.4. The composition of these simulations is given in Table 3. Each simulation was 120 years long, starting from the year 1980 and continuing until the end of the 21st century. We copied the model setup from paper **II**, but restricted our experiment for the Hyytiälä and Sodankylä sites. We used the JSBACH default stomatal conductance formulation and extracted a set of 100 parameter vectors from the corresponding APIS calibrations to represent the JSBACH model uncertainty. The JSBACH model was run with all combinations of parameters and climate model inputs under different RCP scenarios. We also compared our results to those from another model, called PREBAS, which additionally includes forest management actions.

Model	Par	Clim	RCP	Manag	Sites	Total
JSBACH	100	4	2	-	2	1600
PREBAS	100	4	3	2	2	4800

Table 3: Composition of JSBACH and PREBAS model simulations: number of parameter vectors (Par), climate models (Clim), RCP scenarios (RCP), management actions (Manag) and sites as well as the total number of 120-year long simulations.

The JSBACH model produced daily values for a selected set of model output variables. We used the output variables to calculate annual values for given ecosystem indicators of climate change. We produced 30-year averages of these ecosystem indicators to reduce

year-to-year variation. The indicators were calculated for four time periods: 1980-2009 (reference), 2010-2039 (interim), 2040-2069 (mid-century) and 2070-2099 (future). The averaged indicators were also separated into four groups consisting of carbon-, water-, snow- and growing season-related variables.

The variation in the 30-year averaged indicator values was examined with canonical correlation analysis (CCA; Hotelling and Pabst, 1936). This analysis was done separately for all indicators and each indicator group, for each period and for both sites. The results were presented via the use of redundancy indices (Stewart and Love, 1968; Weiss, 1972).

4.4 Essential methods applied in calibration analysis

We have utilised several statistical methods to analyse the different model calibrations. These methods are listed in Table 4 and we will next present some of them in more detail. In addition to the methods in Table 4, we have used many standard measures to assess the model calibration and output variables, e.g. (normalised) root mean squared error estimates, bias, Pearson correlation, regression lines and coefficient of determination.

Method	Abb.	Section	Paper	Reference
Gelman-Rubin diagnostics	$\hat{R}$	4.4.1	**II**	Gelman and Rubin (1992)
Principal component an.	PCA	4.4.2	**I**	Pearson (1901)
Canonical correlation an.	CCA	4.4.3	**III**	Hotelling and Pabst (1936)
Kernel density estimate	KDE		**I-III**	Rosenblatt (1956)
Mann-Kendall trend test			**III**	Mann (1945)
Parameter effectiveness			**I-II**	
Parameter stability	δ		**II**	

Table 4: List of statistical methods used to analyse the calibrations with main references. Selected central methods are introduced in the given sections.

4.4.1 Gelman-Rubin convergence diagnostics

The Gelman-Rubin convergence diagnostics (Gelman and Rubin, 1992) is routinely used to show that multiple MCMC chains have converged to the same target. It is commonly referred to as "a proof of convergence", which it is not – the diagnostics can merely verify that the chains, so far, appear to be converging to the same target.

In essence, the diagnostics compares the variance of each parameter within individual chains (W) to the parameter variance between all the chains (B). W is merely the

average of the individual chain sample variances. B is the sample variance of the chain-specific means, multiplied by the number of draws per chain. Usually, only a subset of all parameter draws are used as the initial-state bias is reduced by discarding some portion of draws from the beginning. The marginal posterior variance of the parameter can be expressed with an unbiased estimator, where n is the number of draws from each chain:

$$\hat{V} = \frac{n-1}{n}W + \frac{1}{n}B. \tag{4.1}$$

The Gelman-Rubin diagnostics states that if the potential scale reduction factor $\hat{R} \approx 1$ for all parameters, the chains have fairly reliably converged to the same target. The potential scale reduction factor is defined as follows:

$$\hat{R} = \sqrt{k\frac{\hat{V}}{W}}. \tag{4.2}$$

The original Gelman-Rubin diagnostics used $k = 1$, but this did not fully account for sampling variability. Brooks and Gelman (1998) corrected this by introducing the factor $k = \frac{d+3}{d+1}$, where d is the estimated degrees of freedom (for the Student t-distribution).

4.4.2 Principal component analysis (PCA)

Principal component analysis (PCA; Pearson, 1901) was used in paper **I** to examine which parameters contribute most to the variation in the calibration process. This was done by examining the posterior distributions from the chains produced by AM. We calculated the parameter covariance matrix ($\boldsymbol{C}$) from the chain history. The covariance matrix undergoes eigenvalue decomposition, where the eigenvectors ($\boldsymbol{v}_i$), eigenvalues (λ_i) and factors ($a_j \in \mathbb{R}$) satisfy the equation:

$$\boldsymbol{C}\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i, \qquad \text{where} \qquad \boldsymbol{v}_i = \sum_j a_j \boldsymbol{\theta}^j \qquad \text{and} \qquad \sum_j a_j^2 = 1. \tag{4.3}$$

Each eigenvector ($\boldsymbol{v}_i$) above is a linear combination of the individual parameter vectors ($\boldsymbol{\theta}^j$). The first eigenvector is called the first principal component and it is constructed to account for the largest variance within the given data. The succeeding eigenvectors are similarly formed but they also required to be orthogonal to the preceding components. The eigenvalue decomposition is unique with mild assumptions, as the covariance matrix is both symmetric and positive semi-definite. The eigenvectors form a new orthogonal basis that optimally represents the variance in the AM chains.

4.4.3 Canonical correlation analysis (CCA) and redundancy indices

Canonical correlation analysis (CCA; Hotelling and Pabst, 1936) is a multivariate extension of correlation analysis. The approach is similar to PCA, but instead of examining variation within one dataset, CCA aims to maximise (linear) correlations between two sets of variables. Our datasets consist of combinations of input f_i and output variables e_j, which represent the uncertainty factors (X) and periodically averaged ecosystem indicators (Y), respectively.

$X \in \mathbb{R}^{s \times n_f}$ and $Y \in \mathbb{R}^{s \times n_e}$ can be represented by matrices, where the row-vectors correspond to the site-specific simulations in Table 3. Here $s = 800$ is the number of simulations, $n_f = 3$ is the number of uncertainty factors and $n_e = 15$ is the number of ecosystem indicators. Each factor $f_i, i \in \{1, ..., n_f\}$, or indicator $e_j, j \in \{1, ..., n_e\}$, can be interpreted as a column-vector of X or Y, respectively. Similarly to eigenvectors in Eq. (4.3), we construct linear combinations of the input factors ($CV_{X_1} = Xa, a \in \mathbb{R}^{n_f}$) and output variables ($CV_{Y_1} = Yb, b \in \mathbb{R}^{n_e}$). The vectors $CV_{X_1}, CV_{Y_1} \in \mathbb{R}^s$ form the first pair of canonical variates (denoted by the subscript 1) as we choose a, b as to maximise the canonical correlation:

$$Rc_1 = \mathrm{corr}(CV_{X_1}, CV_{Y_1}). \tag{4.4}$$

The second pair of canonical variates is formed similarly with an additional requirement that they are uncorrelated with the first pair (and so forth). This results in multiple (orthogonal) canonical pairs (CV_{X_k}, CV_{Y_k}) with diminishing canonical correlations (Rc_k). Here $k \in \{1, ...n_k\}$ and $n_k = 3$ is the number of canonical pairs used in paper **III**.

The simple linear correlations between an independent variable and its canonical variate (f_i and CV_{X_k} or e_j and CV_{Y_k}) are called canonical loadings (CL_{ik}, CL_{jk}). Similarly, the correlations between an independent variable and its opposite canonical variate (f_i and CV_{Y_k} or e_j and CV_{X_k}) are called canonical cross loadings (CcL_{ik}, CcL_{jk}). To summarise the CCA results via the use of a redundancy index (Rd), we need the canonical loadings of the ecosystem indicators (CL_{jk}) and canonical cross loadings of the uncertainty factors (CcL_{ik}).

The redundancy index expresses the amount of variance in a set of variables (ecosystem indicators) explained by another set of variables (uncertainty factors) (Stewart and Love, 1968; Weiss, 1972; van den Wollenberg, 1977). The indices in paper **III** are calculated for all ecosystem indicators and for specific indicator subsets. Here we only present the calculations for the whole set.

$$Rd_{Y_k} = \frac{1}{n_e} \sum_{j=1}^{n_e} (CL_{jk}^2) Rc_k^2. \tag{4.5}$$

The redundancy (Rd_{Y_k}) is an estimate of the variability of CV_{Y_k}, scaled by the k-th

canonical correlation – Rc_k is the only factor (directly) representing the input variables. The redundancy of the ecosystem indicators (Rd_{Y_k}) is used to calculate the importance of the separate input factors (f_i) on the examined ecosystem indicator uncertainty:

$$RdX_{ik} = Rd_{Y_k} CcL_{ik}^2. \tag{4.6}$$

The redundancy of the ecosystem indicators (Rd_{Y_k}) and the input variable impact (RdX_{ik}) are both calculated separately for each canonical variate. The overall redundancy and the full weight of uncertainty for each factor f_i are derived by summing over the canonical variates. This produces an overall measure of the bi-multivariate covariation of the two sets of variables.

5 Discussion of the main results

In this thesis we have utilised two different Bayesian calibration approaches in the JSBACH model. The AM algorithm was chosen for the first calibration experiment, because it is straightforward to use and we had in-house experience in the application of the algorithm. Instead of AM, we could have used many other MCMC methods, such as Hamiltonian Monte Carlo (HMC; Duane et al., 1987) or the differential evolution adaptive Metropolis (DREAM; Vrugt et al., 2009) algorithm. In paper **II**, we turned towards importance sampling to reduce the need for iterative computing. We selected the novel APIS algorithm (Martino et al., 2015) due to its versatility and capability to estimate the parameter expected values and the normalising constant at runtime.

5.1 AM and APIS as calibration methods

Both AM and APIS were suitable for their respective calibration tasks and they were able to produce stable parameter posterior estimates. The standard approach in estimating the effectiveness of importance sampling is to run comparable calibrations with an MCMC sampler and compare the posteriors. We have not produced such estimates as the AM and APIS calibrations should not be directly compared with one another due to differences in model structure, the number of calibrated parameters and differences in the range of allowed parameter values. However, in Fig. 3 we provide an example of the progression of both methods taken from their respective daily calibrations. The parameters represent different levels of identifiability for APIS, estimated in paper **II** with the parameter effectiveness measure listed in Table 4.

We note that the parameter f_{C3} converges to a different value between the two calibrations. This is due to differences in the model structure, but the modelling approaches, the number of parameters and the amount of sites all play a part in the algorithm convergence. We also note that the actual AM simulations were run considerably longer than reported in Table 2, as we wanted to examine the stability of the posterior estimates. This is also the reason for the lack of any stopping criteria in the simulations (for APIS simulations as well).

The AM and APIS have different philosophies in estimating the target distribution. One of the most essential characteristics is the required number of iterations needed to obtain reliable estimates. It is significantly smaller for importance sampling. In Fig. 3 APIS uses 100 consecutive draws per IS sampler (40 samplers) and AM utilises 40,000 draws for each of its eight chains. We note this here only to highlight the differences in the methods, not to proclaim the supremacy of either algorithm. The APIS algorithm was also modified to include adaptation of the proposal shape parameters, an approach that mitigates the use of unconventional likelihoods. Furthermore, it could be beneficial

to use only the latter half of simulated values in the calculation of the APIS global parameter expected values in order to reduce any initial state bias.

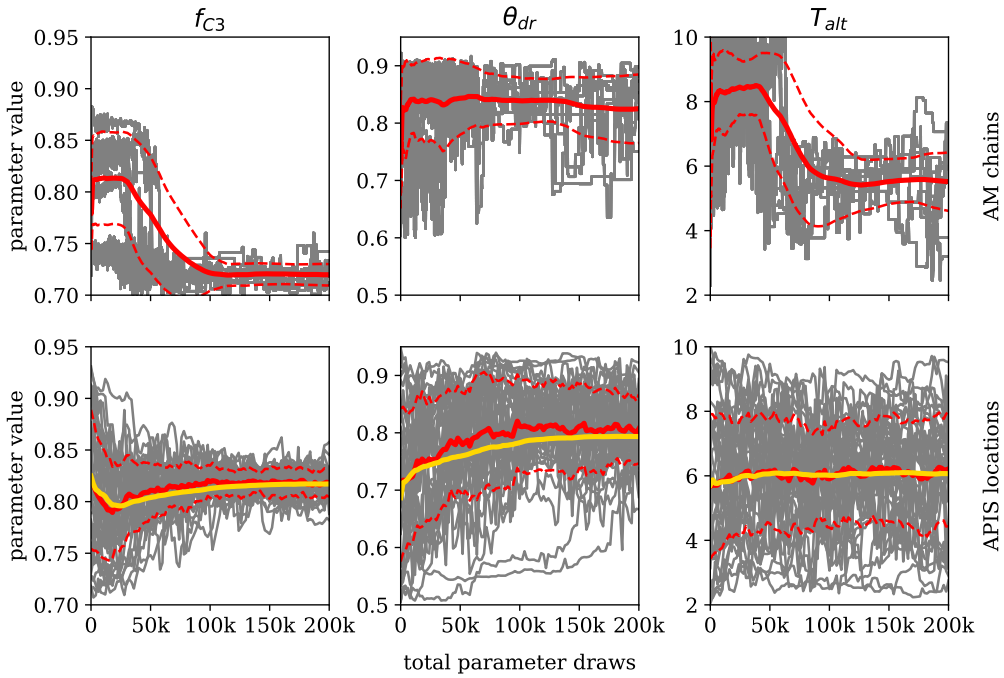

Figure 3: Example of the progression of AM and APIS parameter estimates. AM utilises eight parallel chains and APIS uses 40 independent IS samplers. The red lines are the parameter mean and one standard deviation – for AM this is calculated from the latter half of all previous samples and for APIS it is the current estimate. The yellow line for APIS is the global mean estimate. The parameters are: leaf internal $C0_2$ concentration as a fraction of ambient CO_2 (f_{C3}), fraction of volumetric soil water content above which fast drainage occurs (θ_{dr}) and cut-off value in heat sum to determine spring event (T_{alt}).

5.2 Assessment of model differences between the calibrations

The initial calibration concept in paper **I** allows for site-specific tuning of LAI. Unfortunately, this approach can be applied only for a single site at a time. The method is not as intrusive on the model as readjusting the site level vegetation fractions, applied in paper **II**. The latter approach enables simultaneous models run on multiple sites, but may result in unwanted consequences as e.g. rainfall is made available for plants based on the vegetative fractions. This is in part one of the reasons for poor model performance in paper **II** on sites with low LAI. On the other hand, the model default parameterisation worked roughly as badly for the low LAI sites and the approach was successful for the other sites in the study.

All of the stomatal conductance formulations produced reasonable ET and GPP estimates in paper **II**. Prominent conclusions should not be made solely on the basis of the single drought event for Hyytiälä. The stomatal conductance-specific calibrations resulted in varied parameter estimates, which highlights the importance of more detailed drought processes. The experimental water stress factor (Egea et al., 2011) improved the model performance during the drought. As we adapted only one of three suggested drought pathways in Egea et al. (2011), this presents a multitude of possibilities in future model development. However, calibrating multiple parameters that affect the same process is not straightforward, especially with a limited amount of suitable calibration data.

The delayed effect of temperature for photosynthesis (Mäkelä et al., 2004; Kolari et al., 2007) was highly effective in correcting the springtime GPP increase. Likewise, the inclusion of the multilayer soil moisture scheme (Hagemann and Stacke, 2015) has greatly improved the modelled water cycle. This is not clear from the examined calibration fluxes in paper **II**, but it manifests as more reliable soil moisture values and drought indicators (Gao et al., 2016) and water use efficiency (Gao et al., 2017) estimates that are the basis for some of the ecosystem indicators in paper **III**.

Representation of the soil water cycle could be further enhanced by introducing a more detailed representation of soil configuration, such as vertically varying soil properties. The current version of JSBACH does not support this as the whole soil column is assumed to be structurally homogeneous. Detailed soil fluid flow has been examined by Mattila et al. (2016), who emphasise the importance of pore space on multiple scales in understanding soil water transport capabilities. Even introducing pore sizes, which could vary between the separate layers in the multilayer soil moisture scheme by Hagemann and Stacke (2015), could lead to considerable advances in modelling the soil water cycle and e.g. extreme drought events.

5.3 Reliability of the 21st century ecosystem indicators

The premise of the 21st century ecosystem indicator simulations lies on well-established CMIP5 pathways that have been bias corrected for Finland (Lehtonen et al., 2016; Ruosteenoja et al., 2016) and on model parameterisations that reproduce adequately the current climate conditions. The examined ecosystem indicators also reflect the important processes related to the calibration variables.

The simulations underline the importance of forest management actions as they had the greatest influence on the simulated ecosystem indicator uncertainty, even though the actual management actions were relatively similar. This perceived impact is underscored by the lack of management in many land-surface components of climate models. Similar uncertainty was achieved by the RCP scenarios towards the end of the century,

while climate model uncertainty remains roughly constant throughout the simulations – likely partly due to internal variability of the climate system (Knutti and Sedláček, 2012). Uncertainty related to model parameters was the least influential of the examined uncertainty sources.

The main analysis method used to assess the uncertainties was CCA. The method has caveats as the analysis is only able to account for linear dependencies (Hotelling and Pabst, 1936) between the input and output variables. Hence, any nonlinear and adverse effects are not fully incorporated in the reported uncertainty estimates. This is mostly encountered with the combined climate model and parameter uncertainty that have varied responses to the different stimuli. Especially the parameter uncertainty may be underestimated, which can manifest as e.g. increased drought occurrence.

6 Concluding remarks

One of the most important things in model calibration is to choose the right tool for the problem. If the only aim is to optimise the model parameters, a gradient-based method is the likely candidate. MCMC samplers are currently the most commonly used approaches when more detailed information is required about the parameter posterior distributions. As the computational power increases, namely the number of simultaneous processes, importance samplers can provide an attractive alternative for estimating complex target distributions. This is complemented by the APIS sampler's ability to utilise a mixture of different proposal densities to sample specific regions of the target.

Model calibration is a continuing effort to improve model performance and some calibration should always be made when introducing new model components or processes. Calibration is not done solely for the purpose of optimising parameters, but also to explore caveats in the modelling approach, which is useful when contemplating future model developments.

The APIS calibration provided a parameter set that was used in the 21st century ecosystem indicator uncertainty simulations. These simulations provide assessments of relative uncertainties but can also be used to show where more detailed efforts are required to fully capture the variations in the indicator values. The CCA and redundancy indices were extremely capable in capturing the linear dependencies between the input and output component and their usage should be encouraged. These approaches are also viable in estimating the new CMIP6 model simulations and could be incorporated as evaluation metrics in e.g. ESMValTool (Eyring et al., 2016, 2019).

References

Aurela, M., Lohila, A., Tuovinen, J., Hatakka, J., Penttilä, T., and Laurila, T.: Carbon dioxide and energy flux measurements in four northern-boreal ecosystems at Pallas, Boreal Environ. Res., 20, 455–473, URL http://www.borenv.net/BER/pdfs/ber20/ber20-455.pdf, 2015.

Baldocchi, D., Hicks, B., and Meyers, T.: Measuring biosphere-atmosphere exchanges of biologically related gases with micrometeorological methods, Ecology, 69, 1331–1340, https://doi.org/10.2307/1941631, 1988.

Ball, J., Woodrow, I., and Berry, J.: A Model Predicting Stomatal Conductance and its Contribution to the Control of Photosynthesis Under Different Environmental Conditions, Springer, Progress in Photosynthesis Research (edited by Biggins, J.), 221–224, https://doi.org/10.1007/978-94-017-0519-6_48, 1987.

Boé, J. and Terray, L.: Uncertainties in summer evapotranspiration changes over Europe and implications for regional climate change, Geophys. Res. Lett., 35, L05 702, https://doi.org/10.1029/2007GL032417, 2008.

Brooks, S. P. and Gelman, A.: General Methods for Monitoring Convergence of Iterative Simulations, J. Comput. Graph. Stat., 7, 434–455, https://doi.org/10.1080/10618600.1998.10474787, 1998.

Chen, J., Govind, A., Sonnentag, O., Zhang, Y., Barr, A., and Amiro, B.: Leaf area index measurements at Fluxnet Canada forest sites, Agr. Forest Meteorol., 140, 257–268, https://doi.org/10.1016/j.agrformet.2006.08.005, 2006.

Cornuet, J.-M., Marin, J.-M., Mira, A., and Robert, C.: Adaptive Multiple Importance Sampling, Scand. J. Stat., 39, 798–812, https://doi.org/10.1111/j.1467-9469.2011.00756.x, 2012.

Duane, S., Kennedy, A., Pendleton, B., and Roweth, D.: Hybrid Monte Carlo, Phys. Lett. B, 195, 216–222, https://doi.org/10.1016/0370-2693(87)91197-X, 1987.

Egea, G., Verhoef, A., and Vidale, P.: Towards an improved and more flexible representation of water stress in coupled photosynthesis–stomatal conductance models, Agric. Forest Meteorol., 151, 1370–1384, https://doi.org/10.1016/j.agrformet.2011.05.019, 2011.

Eyring, V., B. S. M. G. A. S. C. A. S. B. S. R. J. and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/gmd-9-1937-2016, 2016.

Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I., Davin, E. L., Deser, C., Ehbrecht, C., Friedlingstein, P., Gleckler, P., Gottschaldt, K.-D., Hagemann, S., Juckes, M., Kindermann, S., Krasting, J., Kunert, D., Levine, R., Loew, A., Mäkelä, J., Martin, G., Mason, E., Phillips, A. S., Read, S., Rio, C., Roehrig, R., Senftleben, D., Sterl, A., van Ulft, L. H., Walton, J., Wang, S., and Williams, K. D.: ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP, Geoscientific Model Development, 9, 1747–1802, https://doi.org/10.5194/gmd-9-1747-2016, 2016.

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., Collins, W. D., Gier, B. K., Hall, A. D., Hoffman, F. M., Hurtt, G. C., Jahn, A., Jones, C. D., Klein, S. A., Krasting, J. P., Kwiatkowski, L., Lorenz, R., Maloney, E., Meehl, G. A., Pendergrass, A. G., Pincus, R., Ruane, A. C., Russell, J. L., Sanderson, B. M., Santer, B. D., Sherwood, S. C., Simpson, I. R., Stouffer, R. J., and Williamson, M. S.: Taking climate model evaluation to the next level, Nat. Clim. Change, 9, 102–110, https://doi.org/10.1038/s41558-018-0355-y, 2019.

Farquhar, G., Caemmerer von, S., and Berry, J.: A Biochemical Model of Photosynthetic CO_2 Assimilation in Leaves of C3 species, Planta, 149, 78–90, https://doi.org/10.1007/BF00386231, 1980.

Friend, A. and Kiang, N.: Land surface model development for the GISS GCM: Effects of improved canopy physiology on simulated climate, J. Climate, 18, 2883–2902, https://doi.org/10.1175/JCLI3425.1, 2005.

Gao, Y.: Interactions between land surface, forests and climate: regional modelling studies in the boreal zone, Finnish Meteorological Institute Contributions, 124, 1–146, URL http://urn.fi/URN:ISBN:978-951-697-894-2, 2016.

Gao, Y., Markkanen, T., Thum, T., Aurela, M., Lohila, A., Mammarella, I., Kämäräinen, M., Hagemann, S., and Aalto, T.: Assessing various drought indicators in representing summer drought in boreal forests in Finland, Hydrol. Earth Syst. Sci., 20, 175–191, https://doi.org/10.5194/hess-20-175-2016, 2016.

Gao, Y., Markkanen, T., Aurela, M., Mammarella, I., Thum, T., Tsuruta, A., Yang, H., and Aalto, T.: Response of water use efficiency to summer drought in boreal Scots pine forests in Finland, Biogeosciences, 14, 4409–4422, https://doi.org/10.5194/bg-14-4409-2017, 2017.

Gelman, A. and Rubin, D.: Inference from Iterative Simulation Using Multiple Sequences, Statist. Sci., 7, 457–472, https://doi.org/10.1214/ss/1177011136, 1992.

Haario, H., Saksman, E., and Tamminen, J.: An adaptive Metropolis algorithm, Bernoulli, 7, 223–242, URL http://projecteuclid.org/euclid.bj/1080222083, 2001.

Hagemann, S. and Stacke, T.: Impact of the soil hydrology scheme on simulated soil moisture memory, Clim. Dynam., 44, 1731–1750, https://doi.org/10.1007/s00382-014-2221-6, 2015.

Hastings, W. K.: Monte Carlo sampling methods using Markov chains and their applications, Biometrika, 57, 97–109, https://doi.org/10.1093/biomet/57.1.97, 1970.

Hotelling, H. and Pabst, M. R.: Rank Correlation and Tests of Significance Involving No Assumption of Normality, The Annals of Mathematical Statistics, pp. 29–43, URL http://www.jstor.org/stable/2957508, 1936.

Ikawa, H., Nakai, T., Busey, R., Kim, Y., Kobayashi, H., Nagai, S., Ueyama, M., Saito, K., Nagano, H., Suzuki, R., and Hinzman, L.: Understory CO_2, sensible heat, and latent heat fluxes in a black spruce forest in interior Alaska, Agr. Forest Meteorol., 214–215, 80–90, https://doi.org/10.1016/j.agrformet.2015.08.247, 2015.

IPCC: Summary for Policymakers, Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, p. 151, [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)] IPCC, Geneva, Switzerland, URL https://www.ipcc.ch/report/ar5/syr/, 2014.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J., Heimann, M., Hooker-Stroud, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E., Köstler, C., Mathieu, P.-P., Pinty, B., Reick, C., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., Terwisscha van Scheltinga, A., Vossbeck, M., Widmann, H., and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, J. Geophys. Res-Biogeo., 118, 1414–1426, https://doi.org/10.1002/jgrg.20118, 2013.

Kelliher, F., Lloyd, J., Arneth, A., Byers, J., McSeveny, T., Milukova, I., Grigoriev, S., Panfyorov, M., Sogatchev, A., Varlargin, A., Ziegler, W., Bauer, G., and Schulze, E.-D.: Evaporation from a central Siberian pine forest, J. Hydrol., 205, 279–296, https://doi.org/10.1016/S0022-1694(98)00082-1, 1998.

Knauer, J., Werner, C., and Zaehle, A.: Evaluating stomatal models and their atmospheric drought response in a land surface scheme: A multibiome analysis, J. Geophys. Res-Biogeo., 120, 1894–1911, https://doi.org/10.1002/2015JG003114, 2015.

Knorr, W.: Satellite Remote Sensing and Modelling of the Global CO_2 Exchange of Land Vegetation: A Synthesis Study, Max-Planck-Institut für Meteorologie Examensarbeit, 49, 1894–1911, URL http://quest.bris.ac.uk/publications/knorr/thesis/thesis.html, 1997.

Knorr, W.: Annual and interannual CO2 exchanges of the terrestrial biosphere: process-based simulations and uncertainties, Global Ecol. Biogeogr., 9, 225–252, https://doi.org/10.1046/j.1365-2699.2000.00159.x, 2000.

Knutti, R. and Sedláček, J.: Robustness and uncertainties in the new CMIP5 climate model projections, Nat. Clim. Change, 3, 369–373, https://doi.org/10.1038/nclimate1716, 2012.

Kolari, P., Lappalainen, H., Hänninen, H., and Hari, P.: Relationship between temperature and the seasonal course of photosynthesis in Scots pine at northern timberline and in southern boreal zone, Tellus B, 59, 542–552, https://doi.org/10.1111/j.1600-0889.2007.00262.x, 2007.

Kolari, P., Kulmala, L., Pumpanen, J., Launiainen, S., Ilvesniemi, H., Hari, P., and Nikinmaa, E.: CO_2 exchange and component CO_2 fluxes of a boreal Scots pine forest, Boreal Environ. Res., 14, 761–783, URL http://www.borenv.net/BER/pdfs/ber14/ber14-761.pdf, 2009.

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., and Richardson, A.: Constraining a global ecosystem model with multi-site eddy-covariance data, Biogeosciences, 9, 3757–3776, https://doi.org/10.5194/bg-9-3757-2012, 2012.

Launiainen, S., Katul, G., Kolari, P., Lindroth, A., Lohila, A., Aurela, M., Varlagin, A., Grelle, A., and Vesala, T.: Do the energy fluxes and surface conductance of boreal coniferous forests in Europe scale with leaf area?, Glob. Change Biol., 22, 4096–4113, https://doi.org/10.1111/gcb.13497, 2016.

Lehtonen, I., Kämäräinen, M., Gregow, H., Venäläinen, A., and Peltola, H.: Heavy snow loads in Finnish forests respond regionally asymmetrically to projected climate change, Natural Hazards and Earth System Sciences, 16, 2259–2271, https://doi.org/10.5194/nhess-16-2259-2016, 2016.

Leuning, R.: A critical appraisal of a combined stomatal-photosynthesis model for C_3 plants, Plant Cell Environ., 18, 339–355, https://doi.org/10.1111/j.1365-3040.1995.tb00370.x, 1995.

Mäkelä, A., Hari, P., Berninger, F., Hänninen, H., and Nikinmaa, E.: Acclimation of photosynthetic capacity in Scots pine to the annual cycle of temperature, Tree Physiol., 24, 369–376, https://doi.org/10.1093/treephys/24.4.369, 2004.

Mann, H. B.: Nonparametric Tests Against Trend, Econometrica, 13, 245–259, https://doi.org/10.2307/1907187, 1945.

Martino, L., Elvira, V., Luengo, D., and Corander, J.: An Adaptive Population Importance Sampler: Learning From Uncertainty., IEEE Transactions on Signal Processing, 63, 4422–4437, https://doi.org/10.1109/TSP.2015.2440215, 2015.

Mattila, K., Puurtinen, T., Hyväluoma, J., Surmas, R., Myllys, M., Turpeinen, T., Robertsén, F., Westerholm, J., and Timonen, J.: A prospect for computing in porous materials research: Very large

fluid flow simulations, J. Comput. Sci., 12, 1877–7503, https://doi.org/10.1016/j.jocs.2015.11.013, 2016.

Medlyn, B., Duursma, R., Eamus, D., Ellsworth, D., Prentice, I., Barton, C., Crous, K., De Angelis, P., Freeman, M., and Wingate, L.: Reconciling the optimal and empirical approaches to modelling stomatal conductance, Glob. Change Biol., 17, 2134–2144, https://doi.org/10.1111/j.1365-2486.2010.02375.x, 2011.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E.: Equations of State Calculations by Fast Computing Machines, J. ofChem. Phys., 21, 1087–1091, https://doi.org/10.1063/1.1699114, 1953.

Minunno, F., Peltoniemi, M., Härkönen, S., Kalliokoski, T., Makinen, H., and Mäkelä, A.: Bayesian calibration of a carbon balance model PREBAS using data from permanent growth experiments and national forest inventory, Forest Ecol. Manag., 440, 208–257, https://doi.org/10.1016/j.foreco.2019.02.041, 2019.

Mueller, B. and Seneviratne, S.: Systematic land climate and evapotranspiration biases in CMIP5 simulations, Geophys. Res. Lett., 41, 128–134, https://doi.org/10.1002/2013GL058055, 2014.

Owen, A. and Yi, Z.: Safe and Effective Importance Sampling, J. Am. Stat. Assoc., 95, 135–143, https://doi.org/10.2307/2669533, 2000.

Parzen, E.: On Estimation of a Probability Density Function and Mode, Ann. Math. Statist., 33, 1065–1076, https://doi.org/10.1214/aoms/1177704472, 1962.

Pearson, K.: LIII. On lines and planes of closest fit to systems of points in space, Philos. Mag., 2, 559–572, https://doi.org/10.1080/14786440109462720, 1901.

Peltoniemi, M., Markkanen, T., Härkönen, S., Muukkonen, P., Thum, T., Aalto, T., and Mäkelä, A.: Consistent estimates of gross primary production of Finnish forests – comparison of estimates of two process models, Boreal Environ. Res., 20, 196–212, URL http://www.borenv.net/BER/pdfs/ber20/ber20-196.pdf, 2015a.

Peltoniemi, M., Pulkkinen, M., Aurela, M., Pumpanen, J., Kolari, P., and Mäkelä, A.: A semi-empirical model of boreal-forest gross primary production, evapotranspiration, and soil water – calibration and sensitivity analysis, Boreal Environ. Res., 20, 151–171, URL http://www.borenv.net/BER/pdfs/ber20/ber20-151.pdf, 2015b.

Raddatz, T., Reick, C., Korr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K.-G., Wetzel, P., and Jungclau, J.: Will the tropical land biosphere dominate the climate–carbon cycle feedback during the twenty-first century?, Clim. Dynam., 29, 565–574, https://doi.org/10.1007/s00382-007-0247-8, 2007.

Rannik, U., Peltola, O., and Mammarella, I.: Random uncertainties of flux measurements by the eddy covariance technique, Atmos. Meas. Tech. Discuss., in review, 1–31, https://doi.org/10.5194/amt-2016-31, 2016.

Rantala, M., Leskinen, L., Hujala, T., and Kurttila, M.: Arvio METSO-ohjelman yhteistoimintaverkostohankkeiden vaikuttavuudesta ja kehittämistarpeista, Working Papers of the Finnish Forest Research Institute, 202, URL www.metla.fi/julkaisut/workingpapers/2011/mwp202.htm, 2011.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F.,

Vesala, T., Yakir, D., and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, Glob. Change Biol., 11, 1424–1439, https://doi.org/10.1111/j.1365-2486.2005.001002.x, 2005.

Reick, C., Raddatz, T., Brovkin, V., and Gayler, V.: Representation of natural and anthropogenic land cover change in MPI-ESM, J. Adv. Model. Earth Syst., 5, 1–24, https://doi.org/10.1002/jame.20022, 2013.

Richardson, A., Hollinger, D., Burba, G., Davis, K., Flanagan, L., Katul, G., Munger, J., Ricciutio, D., Stoy, P., Suyker, A., Verma, S., and Wofsy, S.: A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes, Agr. Forest Meteorol., 136, 1–18, https://doi.org/10.1016/j.agrformet.2006.01.007, 2006.

Richardson, A., Mahecha, M., Falge, E., Kattge, J., Moffat, A., Papale, D., Reichstein, M., Stauch, V., Braswell, B., Churkina, G., Kruijt, B., and Hollinger, D.: Statistical properties of random CO2 flux measurement uncertainty inferred from model residuals, Agr. Forest Meteorol., 148, 38–50, https://doi.org/10.1016/j.agrformet.2007.09.001, 2008.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM5. PART I: Model description, Max Planck Institute for Meteorology Report, 349, 1–127, URL http://www.mpimet.mpg.de/fileadmin/publikationen/Reports/max_scirep_349.pdf, 2003.

Rosenblatt, M.: Remarks on Some Nonparametric Estimates of a Density Function, Ann. Math. Statist., 27, 832–837, https://doi.org/10.1214/aoms/1177728190, 1956.

Ruosteenoja, K., Jylhä, K., and Kämäräinena, M.: Climate Projections for Finland Under the RCP Forcing Scenarios, Geophysica, 51, 17–50, 2016.

Sellers, P.: Canopy reflectance, photosynthesis and transpiration, Int. J. Remote Sens., 6, 1335–1372, https://doi.org/10.1080/01431168508948283, 1985.

Stewart, D. and Love, W.: A general canonical correlation index, Psychol. Bull., 70, 160–163, https://doi.org/10.1037/h0026143, 1968.

Thum, T., Aalto, T., Laurila, T., Aurela, M., Kolari, P., and Hari, P.: Parametrization of two photosynthesis models at the canopy scale in northern boreal Scots pine forest, Tellus, 59B, 874–890, https://doi.org/10.1111/j.1600-0889.2007.00305.x, 2007.

Valentine, H. T. and Mäkelä, A.: Bridging process-based and empirical approaches to modeling tree growth, Tree Physiol., 25, 769–779, https://doi.org/10.1093/treephys/25.7.769, 2005.

van den Wollenberg, A.: Redundancy analysis an alternative for canonical correlation analysis, Psychometrika, 42, 207–219, https://doi.org/10.1007/BF02294050, 1977.

Veach, E. and Guibas, L.: Optimally Combining Sampling Techniques for Monte Carlo Rendering, SIGGRAPH 1995 Proceedings, pp. 419–428, https://doi.org/10.1145/218380.218498, 1995.

Vrugt, J., ter Braak, C., Diks, C., Robinson, B., Hyman, J., and Higdon, D.: Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, Int. J. Nonlin. Sci. Num., 10, 273–290, https://doi.org/10.1515/IJNSNS.2009.10.3.273, 2009.

Weiss, D. J.: Canonical correlation analysis in counseling psychology research, J. Couns. Psychol., 19, 241–252, https://doi.org/10.1037/h0032675, 1972.

I

Paper I

Constraining ecosystem model with adaptive Metropolis algorithm using boreal forest site eddy covariance measurements

Mäkelä, J., Susiluoto, J., Markkanen, T., Aurela, M., Järvinen, H., Mammarella, I., Hagemann, S., and Aalto, T.

Nonlin. Processes Geophys., 23, 447–465,
doi:10.5194/npg-23-447-2016, 2016.

©Authors 2016.
Reprinted under Creative Commons Attribution 3.0 License.

Nonlin. Processes Geophys., 23, 447–465, 2016
www.nonlin-processes-geophys.net/23/447/2016/
doi:10.5194/npg-23-447-2016
© Author(s) 2016. CC Attribution 3.0 License.

Nonlinear Processes in Geophysics
Open Access
EGU

Constraining ecosystem model with adaptive Metropolis algorithm using boreal forest site eddy covariance measurements

Jarmo Mäkelä[1], Jouni Susiluoto[1], Tiina Markkanen[1], Mika Aurela[1], Heikki Järvinen[2], Ivan Mammarella[2], Stefan Hagemann[3], and Tuula Aalto[1]

[1]Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland
[2]Department of Physics, P.O. Box 48, University of Helsinki, 00014 Helsinki, Finland
[3]Max Planck Institute for Meteorology, Bundesstraße 53, 20146 Hamburg, Germany

Correspondence to: Jarmo Mäkelä (jarmo.makela@fmi.fi)

Received: 29 February 2016 – Published in Nonlin. Processes Geophys. Discuss.: 11 April 2016
Revised: 17 October 2016 – Accepted: 21 November 2016 – Published: 9 December 2016

Abstract. We examined parameter optimisation in the JSBACH (Kaminski et al., 2013; Knorr and Kattge, 2005; Reick et al., 2013) ecosystem model, applied to two boreal forest sites (Hyytiälä and Sodankylä) in Finland. We identified and tested key parameters in soil hydrology and forest water and carbon-exchange-related formulations, and optimised them using the adaptive Metropolis (AM) algorithm for Hyytiälä with a 5-year calibration period (2000–2004) followed by a 4-year validation period (2005–2008). Sodankylä acted as an independent validation site, where optimisations were not made.

The tuning provided estimates for full distribution of possible parameters, along with information about correlation, sensitivity and identifiability. Some parameters were correlated with each other due to a phenomenological connection between carbon uptake and water stress or other connections due to the set-up of the model formulations. The latter holds especially for vegetation phenology parameters. The least identifiable parameters include phenology parameters, parameters connecting relative humidity and soil dryness, and the field capacity of the skin reservoir. These soil parameters were masked by the large contribution from vegetation transpiration.

In addition to leaf area index and the maximum carboxylation rate, the most effective parameters adjusting the gross primary production (GPP) and evapotranspiration (ET) fluxes in seasonal tuning were related to soil wilting point, drainage and moisture stress imposed on vegetation. For daily and half-hourly tunings the most important parameters were the ratio of leaf internal CO_2 concentration to external CO_2 and the parameter connecting relative humidity and soil dryness. Effectively the seasonal tuning transferred water from soil moisture into ET, and daily and half-hourly tunings reversed this process.

The seasonal tuning improved the month-to-month development of GPP and ET, and produced the most stable estimates of water use efficiency. When compared to the seasonal tuning, the daily tuning is worse on the seasonal scale. However, daily parametrisation reproduced the observations for average diurnal cycle best, except for the GPP for Sodankylä validation period, where half-hourly tuned parameters were better. In general, the daily tuning provided the largest reduction in model–data mismatch.

The models response to drought was unaffected by our parametrisations and further studies are needed into enhancing the dry response in JSBACH.

1 Introduction

Inverse modelling of ecosystem model parameters against in situ observations is an established way to tune model parameters (e.g. Scharnagl et al., 2011). As observation sites have their own characteristics, it is necessary to make local site simulations for model evaluation and calibration as they may reveal new insight into model behaviour and guide further development. Model–data fusion has been applied for boreal forest sites by, e.g., Aalto et al. (2004) Peltoniemi et al. (2015b), Thum et al. (2007, 2008) and Wu et al. (2011).

In this study we perform site level parameter optimisation in the JSBACH model (Kaminski et al., 2013; Knorr and Kattge, 2005; Reick et al., 2013). JSBACH is the land surface component of the Earth system model of the Max Planck Institute for Meteorology (MPI-ESM), used to simulate water and carbon storages and fluxes in the soil–vegetation–atmosphere continuum. The water and carbon fluxes are coupled in the model and thus modification of parameters related to one component usually has an effect on the others as well. The optimisation process and the optimised values are also affected by the assimilation frequency and interval in minimising the model–data mismatch. This effect can be studied in numerous ways; e.g. Santaren et al. (2014) varied the length of assimilation interval while Matheny et al. (2014) focused on the diurnal error patterns.

The motivation for this study comes from results showing that CMIP5 model simulations, one of which is MPI-ESM, have systematic evapotranspiration biases over continental areas (Mueller and Seneviratne, 2014). These kinds of biases not only have significant implications for climate change projections (Boé and Terray, 2008) but also have a distinctive behaviour on a regional scale. In addition, a comparative study of the gross primary production (GPP) of Finnish forests (Peltoniemi et al., 2015a) revealed that JSBACH has an insufficient response to water limitation in Finland – it tends to overestimate GPP and evapotranspiration during dry periods. This is especially apparent in the dry year 2006, as JSBACH is unable to transfer the reduced rainfall into lower levels of GPP.

In this study we apply the JSBACH ecosystem model for Hyytiälä (Kolari et al., 2009; Suni et al., 2003) and Sodankylä (Aurela, 2005; Thum et al., 2008) sites. We identify key parameters in soil hydrology and evapotranspiration-related formulations and test their effectiveness with elementary methods. We study the effect of different timescales (seasonal, daily and half-hourly) on the assimilation process and the effect of this on the optimised parameter values. Several optimisations are performed using the adaptive Metropolis (AM) algorithm over a 5-year calibration period (2000–2004) followed by a 4-year validation period (2005–2008).

The goals of this study are to test the applicability of the AM optimisation method for JSBACH and the impact of different temporal resolutions on the optimisation process, and to improve the models response to environmental drivers, focusing on dryness.

2 Materials and methods

2.1 Measurements, sites and instrumentation

In this study we use site level data from two Finnish measurement sites: Hyytiälä (61°51′ N, 24°17′ E; 180 m a.s.l.) and Sodankylä (67°22′ N, 26°38′ E; 179 m a.s.l.). These well-established sites have long continuous measurement data sets representing the southern and northern boreal Finnish coniferous evergreen forests. The data used in this study are available for the scientific community through various databases such as FLUXNET (re3data.org, 2016).

Hyytiälä site is a Finnish Scots pine (*Pinus sylvestris*) forest (Kolari et al., 2009), planted in 1962 after burning and mechanical soil preparation. The soil type in Hyytiälä is Haplic Podzol on glacial till and the site is of medium fertility (Kolari et al., 2009). The forest also has sparse understory of Norway Spruce (*Picea abies*) and scattered deciduous trees. The maximum of measured all-sided leaf area index (LAI) is $6.5\,m^2\,m^{-2}$ for the Scots pine. The carbon dioxide (CO_2) and water vapour (H_2O) fluxes between vegetation and atmosphere have been measured in Hyytiälä continuously since 1997 (Vesala et al., 2005).

The Sodankylä forest, in Sodankylä at the Finnish Meteorological Institute's Arctic Research Centre, is also a Scots pine forest (*Pinus sylvestris*) with maximum measured LAI of $3.6\,m^2\,m^{-2}$ as determined from a forest inventory in early autumn (Thum et al., 2007). The forest on fluvial sandy Podzol has been naturally regenerated after forest fires with tree age ranging approximately from 50 to 100 years. The sparse ground vegetation consists of lichens (73 %), mosses (12 %) and ericaceous shrubs (15 %). The CO_2 and H_2O flux measurements in Sodankylä have been running since 2000 (Aurela, 2005).

The CO_2 and H_2O fluxes were measured by the micrometeorological eddy covariance (EC) method, which provides a direct measurement of the mass and energy exchange between the atmosphere and the biosphere averaged on an ecosystem scale. In the EC method, the flux is obtained as the covariance of the high-frequency (10 Hz) observations of vertical wind speed and the constituent in question (Baldocchi, 2003). The CO_2 fluxes were corrected for the storage change below the measurement height to accurately estimate the net ecosystem CO_2 exchange (NEE). The GPP was derived by subtracting the modelled respiration (R) from the NEE observation ($\mathrm{GPP} = \mathrm{NEE} - R$) utilising standard flux partitioning procedures (Reichstein et al., 2005; Kolari et al., 2009). By using the same parametrisations as in the partitioning, the NEE and GPP time series were gap-filled for comparison with the model results. The daily evapotranspiration (ET) sums were calculated from H_2O flux data that were gap-filled based on the mean diurnal cycles or regressions on available radiative energy.

The EC instrumentation in Hyytiälä consisted of a Solent 1012R3 three-axis sonic anemometer (Gill Instruments Ltd., Lymington, UK) and a LI-6262 closed-path CO_2/H_2O gas analyser (Li-Cor Inc., Lincoln, NE, USA), while in Sodankylä a USA-1 (METEK GmbH, Elmshorn, Germany) anemometer and an LI-7000 (Li-Cor., Inc., Lincoln, NE, USA) closed-path gas analyser were used. The EC fluxes were calculated as half-hourly averages taking into account the required corrections. The measurement systems and the post-processing procedures have been presented in

more detail for Hyytiälä by Kolari et al. (2004) and Mammarella et al. (2009), and for Sodankylä by Aurela (2005) and Aurela et al. (2009).

The measurement error in the EC flux data may be classified into two categories: systematic errors and random errors. The main systematic errors (density fluctuations, high-frequency losses, calibration issues) are mostly corrected for as part of the post-processing of the data, and the random errors tend to dominate the uncertainty of the instantaneous fluxes. The random error is often assumed Gaussian but can be more accurately approximated by a symmetric exponential distribution (Richardson et al., 2006). It increases linearly with the magnitude of the flux, with a standard deviation typically less than 20 % of the flux (Richardson et al., 2008; Rannik et al., 2016).

2.2 The JSBACH model

JSBACH is a process-based ecosystem model and the land surface component of the MPI-ESM. We used JSBACH offline using an observational atmospheric data set to force the model. Implications of this one-way coupling with the atmosphere include lack of feedback from the surface energy balance to the atmosphere; i.e. latent and sensible heat fluxes and surface thermal radiation do not directly affect prescribed air temperature or humidity. Similarly, the feedback of surface to the vertical transfer coefficients within the atmospheric surface layer is missing, as the wind speed that drives mixing is prescribed. Furthermore, since we use site level data (a single grid point), the grid resolution does not affect the results (Tesfa et al., 2014; Singh et al., 2015). We give here a general introduction to JSBACH, whereas a more complete model description can be found in Roeckner et al. (2003).

In JSBACH the land surface is a fractional structure where the land grid cells are divided into tiles representing the most prevalent vegetation classes called plant functional types (PFTs) within each grid cell (Reick et al., 2013). The grid cell is first divided into bare soil and vegetative area which is furthermore fractionally divided into PFTs. The model was set up to effectively use only one tile, coniferous evergreen trees. Henceforth, all model and process descriptions are considered in relation to coniferous evergreen trees and no distinction between PFT-specific and general parameters are made in this study.

Coniferous evergreen trees are characterised by a set of parameters that control vegetation-related biological and physical processes accounting for the land–atmosphere interactions. We made use of expert knowledge to set these parameters for our local sites and verified that they are in line with those presented by Hagemann (2002) and Hagemann and Stacke (2015).

The seasonal development of LAI is regulated by air temperature and soil moisture with a specific maximum LAI as a limiting value. The cycle is driven by a pseudo soil temperature that is a weighted running mean of air temperature. The predictions of phenology are produced by the Logistic Growth Phenology (LoGro-P) model of JSBACH.

Photosynthesis is described by the biochemical photosynthesis model (Farquhar et al., 1980). Following Kattge et al. (2009), we set the maximum carboxylation rate at 25 °C to 1.9 times the maximum electron transport rate at 25 °C.

The photosynthetic rate is resolved in two steps. First the stomatal conductance under conditions with no water stress is assumed to be controlled by photosynthetic activity (Schulze et al., 1994). Here the leaf internal CO_2 concentration is assumed to be a constant fraction of ambient concentration, which allows for an explicit resolution of the photosynthesis (Knorr, 1997). Then the impact of soil water availability is accounted for by a soil moisture-dependent multiplier that is identical for each canopy layer (Knorr, 1997).

Radiation absorption is estimated by a two-stream approximation within a three-layer canopy (Sellers, 1985). Especially in the sparse canopies, the radiation absorption is affected by clumping of the leaves, which is here taken into account according to the formulation by Knorr (1997).

2.3 The JSBACH model spin-up and runs

Before tuning the JSBACH model, some of the more slowly changing variables (e.g. LAI) need to be equilibrated in order to bring the model into a (semi-)steady state. We achieve this by running the model through a spin-up period generated by looping the measurement interval over itself. During this period the necessary variables are equilibrated and their values become acceptable for the tuning process. At the end of the spin-up a restart file is generated so that the model can be restarted from this state.

We use half-hourly measurements from 1999 to 2008 for Hyytiälä. The spin-up finishes at the end of 1999 and is followed by a calibration period (abbreviated as HC for Hyytiälä calibration) of 2000–2004 and a validation period (HV for Hyytiälä validation) of 2005–2008, including an exceptionally dry summer in 2006. The set-up for Sodankylä is similar but we use measurements from 2000 to 2008, where the spin-up finishes at the end of 2008. The model is then restarted from the start of 2000, but we only examine the Sodankylä validation period (SV) of 2005–2008. The main reason to exclude the Sodankylä calibration period is that essentially we do not calibrate (or tune) the model for Sodankylä and we do not want to appear to do so.

The meteorological data used to drive the climate were air temperature, air pressure, atmospheric CO_2 concentration, precipitation, specific humidity, short- and longwave radiation, potential shortwave radiation and wind speed.

2.4 The parameters

The JSBACH model was modified to fit our custom-built test bed so that all parameters of interest could be read from an

Table 1. Parameter descriptions with references to equations in Appendix A.

Parameter	Units	Class	Description	
Δ_{max}	–	I	Maximum all-sided leaf area index that vegetation can reach.	Eq. (A1)
$V_{C,max}$	◇	I	Farquhar model maximum carboxylation rate at 25 °C of the enzyme Rubisco (coupled with maximum electron transport rate at 25 °C with a factor of 1.9) [◇ = μmol (CO_2) m^{-2} s^{-1}].	Eq. (A8)
veg_{max}	–	I	Fraction of vegetative soil in a grid cell. The rest is bare soil.	–
α_q	–	II	Farquhar model efficiency for photon capture at 25 °C.	Eq. (A11)
c_b	–	II	Adjustment parameter used in stability functions for momentum and heat (Louis, 1979).	–
f_{C3}	–	II	Ratio of C3-plant internal/external CO_2 concentration.	Eq. (A9)
p_{int}	–	II	Fraction of precipitation intercepted by the canopy.	Eq. (A12)
w_{dr}	–	II	Critical fraction of field capacity above which fast drainage occurs for soil water content.	–
w_{hum}	–	II	Fraction depicting relative humidity based on soil dryness.	Eq. (A17)
w_{pwp}	–	II	Fraction of soil moisture at permanent wilting point.	Eq. (A15)
w_{skin}	m	II	Maximum water content of the skin reservoir of bare soil.	–
w_{tsp}	–	II	Fraction of soil moisture above which transpiration is not affected by soil moisture stress.	Eq. (A15)
s^*_{sm}	m	II	Depth for correction of surface temperature for snowmelt.	–
T_{alt}	°C	III	LoGro phenology: alternating temperature. Cut-off temperature used for calculating heat sum to determine the spring event and the number of chill days since the last autumn event.	Eqs. (A2), (A3)
C^*_{decay}	–	III	LoGro phenology: memory loss parameter for chill days.	Eq. (A4)
S_{min}	°C	III	LoGro phenology: minimum value of critical heat sum.	Eq. (A4)
S^*_{range}	°C	III	LoGro phenology: maximal range of critical heat sum.	Eq. (A4)
T_{ps}	°C	III	LoGro phenology: memory loss parameter for calculating pseudo soil temperature.	Eq. (A6)

* These parameters were tested but yielded no or only a minimal response to cost functions and were thus removed from the trial.

external file. We examined 15 parameters (Table 1) that are for convenience separated into three classes. The class I parameters are used differently from those of class II and III – namely, class I parameters are only tuned in the seasonal tuning (explained in detail in Sect. 3.1). Additionally, the only distinction between class II and III parameters is that the latter belong to a specific part of JSBACH called the LoGro-P – there is no difference in how these parameters are used. We also note that the only parameter (of those examined) that can vary from site to site is veg_{max} (the vegetative fraction of a grid cell).

2.5 Parameter sampling

The parameter sampling in this study was done with the AM algorithm. The AM algorithm is an adaptive Markov chain Monte Carlo (MCMC) process described below (it is not strictly Markovian but satisfies the necessary ergodicity requirements). AM is based on the classical Metropolis algorithm, extended with the adaptation of the parameter proposal distribution. Due to the adaptive nature of AM, it does not rely on the choice of the initial proposal distribution. AM is a sampling method that produces estimates of the full distribution of possible parameter values (unlike straightforward optimisation methods), thus enabling, e.g., the study of parameter identifiability, sensitivity and (nonlinear) correlation – this information is paramount to understanding the optimisation process in contrast to merely receiving the optimised parameter values. The rigorous mathematical presentation of the AM algorithm is given in Haario et al. (2001).

The AM algorithm draws samples (sets of parameters) from the parameter space to generate probability distributions for the parameters. The consecutive draws form an MCMC chain. We used the algorithm simultaneously for several independent chains that are parallel adaptations of the algorithmic process (see e.g. Craiu et al., 2009; Solonen et al., 2012) – we take several random starting points and launch the algorithm for each of these simultaneously. The history of all chains is used for updating the proposal covariance matrix that describes how the parameters relate to one another. Our initial proposal covariance matrix had diagonal elements corresponding to 1/200 of the respective parameter's range. The first sample for each chain was chosen at random within this range. The algorithmic process can be described by a few steps:

1. Draw a new sample (x') of the parameter space from the vicinity of the current sample (x) using the current proposal covariance matrix.

2. Calculate the acceptance ratio (a) for the drawn sample. This is the value of a likelihood function (f) that is proportional to the desired probability distribution, at the drawn sample divided by the value at the current sample ($a = f(x')/f(x)$).

3. Accept the new candidate (x') with the probability a (if $a \geq 1$, we always accept). If the candidate was rejected, the current sample (x) is reused as a basis of the next draw and repeated in the chain. Update the covariance matrix and draw a new sample.

We obtain the likelihood function (f) from the cost functions (cf) described below by assuming Gaussian error statistics and setting $f = e^{-\mathrm{cf}}$. In general to estimate the distribution of parameters of any model based on some data, we require some information about the underlying measurement and modelling errors. We treat the JSBACH model as described by the equation $\boldsymbol{y} = M(\boldsymbol{x}, \theta) + \boldsymbol{e}$. Here $\boldsymbol{y}$ are the observations, $\boldsymbol{x}$ is the model state vector, θ are the current parameters and $\boldsymbol{e}$ is the model–data mismatch. Since we only have a robust estimate for the measurement errors and no true error statistics for the model, the full error ($\boldsymbol{e}$) is treated as Gaussian white noise.

The cost function (Eq. 1) used in this study for seasonal tuning is based on summary statistics of GPP and evapotranspiration (ET) along with the maximum of LAI. The cost function (Eq. 1) calculates the relative error in total GPP, ET and growing season maximum of LAI against observations (these are respectively denoted as G_1, E_1 and L_1) and sums them up. Overlined variables refer to the mean value of that variable for a given period (calibration or validation), subscripts denote observation or model results.

$$\mathrm{cf}_1 = \overbrace{\left(\frac{\overline{\mathrm{GPP}}_{\mathrm{mod}} - \overline{\mathrm{GPP}}_{\mathrm{obs}}}{\overline{\mathrm{GPP}}_{\mathrm{obs}}}\right)^2}^{G_1} + \overbrace{\left(\frac{\overline{\mathrm{ET}}_{\mathrm{mod}} - \overline{\mathrm{ET}}_{\mathrm{obs}}}{\overline{\mathrm{ET}}_{\mathrm{obs}}}\right)^2}^{E_1} + \overbrace{\left(\frac{\max(\mathrm{LAI}_{\mathrm{mod}}) - \max(\mathrm{LAI}_{\mathrm{obs}})}{\max(\mathrm{LAI}_{\mathrm{obs}})}\right)^2}^{L_1} \quad (1)$$

The second cost function (Eq. 2) is a slightly modified mean squared error estimate used for daily (cf_2) and half-hourly (cf_3) tuning. With multiple variables there is always the problem of having one variable dominating over the others. Since no true errors were available, it was decided to normalise the residuals using the mean of observations in the cost function (Eq. 2). This way the resulting function is sensitive to changes in both variables – AM is used as a noise-resistant optimiser and sampling is done in the spirit of studying the identifiability and correlations of the parameters. The components are denoted as G_2, E_2 for daily and G_3, E_3 for half-hourly tuning.

$$\mathrm{cf}_{2,3} = \overbrace{\frac{1}{N_{\mathrm{GPP}}}\sum\left(\frac{\mathrm{GPP}_{\mathrm{mod}} - \mathrm{GPP}_{\mathrm{obs}}}{\overline{\mathrm{GPP}}_{\mathrm{obs}}}\right)^2}^{G_{2,3}} + \overbrace{\frac{1}{N_{\mathrm{ET}}}\sum\left(\frac{\mathrm{ET}_{\mathrm{mod}} - \mathrm{ET}_{\mathrm{obs}}}{\overline{\mathrm{ET}}_{\mathrm{obs}}}\right)^2}^{E_{2,3}} \quad (2)$$

As noted previously, JSBACH was used uncoupled from the other components of the full MPI-ESM. This has a tendency to lead to biased results in the model runs as has been recently studied by Dalmonech et al. (2015). Especially in the high latitudes, evapotranspiration can be unrealistic during winter since night-time is longer and temperatures low. In order to improve the credibility of our results, we masked the evapotranspiration values of the coldest periods, and only took into account those from May to September for each year in both cost functions.

2.6 Parameter analysis

The optimised parameter values are taken as the mean values of all chains in the sampling process. In the case that the parameter chains converge to a bound of an a priori prescribed range of allowed values, the maximum a posteriori (MAP) value is used instead. After tuning the model, we analysed different aspects of this process. Class I parameters are excluded from this analysis since they are used to bring the model to an "acceptable initial state"; hence, we regard them as a part of the model initialisation (our motivation is explained in Sect. 3.1).

We calculated the correlations and correlation matrices between different parameters for different tunings using the tested parameter vectors in the AM process. Then we performed a principal component analysis (PCA) on the correlation matrices to get the eigenvectors ($\boldsymbol{v}_{\mathrm{i}}$) and eigenvalues ($e_{\mathrm{i}}$) of the least identifiable parameters in the tuning process with the given data. The PCA transforms the correlation matrix into an orthogonal form where the eigenvector related to the greatest eigenvalue is the least identifiable with the given data. We then calculate the weight ($w_{\mathrm{i}} = \sqrt{\frac{e_{\mathrm{i}}^2}{\sum_i e_{\mathrm{i}}^2}}$) for each component (or vector $\boldsymbol{v}_{\mathrm{i}}$; note that the squared weights sum up to one). We also determine the most dominant parameters for each component ($\boldsymbol{v}_{\mathrm{i}}$) by similarly dividing the length of the vector towards that parameter by the length of the whole vector (weight of vector components).

The information derived with PCA could be extracted by analysing the parameters posterior probability distributions, but PCA yields a simple, straightforward method for the same purpose. The main caveat of the standard PCA method is that it is not applicable to cases with strong nonlinear correlations. Therefore, we also calculate kernel density estimates (KDE) for the parameters to show that there are no nonlinear correlations. The KDE method places a Gaussian distribution (kernels) centred at each parameter of the MCMC chain and then sums these kernels to produce an estimate for the whole distribution. The bandwidth is calculated using the Scott's rule (Scott, 2004).

We also wanted to examine which parameters contributed the most to the change in the cost function values as we switched from one parameter set to another. This was done by calculating the change in the cost function values of the tuned parameter set and a set where one parameter has been reverted to the value the tuning started with (henceforth, the reference values are for seasonal tuning the default values

and for daily and half-hourly tuning the seasonally tuned values). We call this method "relative effectiveness", since we want to analyse the effect of the parameters to the cost function. For each tuned set of parameter values, the relative effectiveness of a parameter is calculated as follows:

1. change one parameter from the set of tuned parameter values to a reference value and calculate the difference in the cost function for the changed set and the tuned set;
2. return the changed parameter to the tuned value and repeat for all parameters (sum up the differences);
3. the relative effectiveness for each parameter is the difference obtained from step 1 divided by the sum from step 2.

The relative effectiveness is similar to a class of methods commonly referred to as the one-at-a-time (OAT) or one-factor-at-a-time (OFAT) methods. These methods are generally used to acquire robust information about model behaviour when one parameter at a time is changed to a new and hopefully better value (e.g. Murphy et al., 2004). The main difference of our method to classical methods such as the Morris OAT (Morris, 1991) is that in such methods the change in values is (usually) random, whereas we have fixed values. Additionally, our point of view is from the optimised parameters to the original state – we have already optimised the parameters (as a group) and merely want some robust and easily comprehensible information about the effect of changes in parameter values to the cost functions. This method does not reveal information about how well the parameters constrain the cost function (e.g. we could have a highly dominating parameter that would optimise to the default value and hence the relative effectiveness would be zero), rather which parameters contribute most to the change in cost function values.

Lastly, we calculate the root mean squared error (RMSE; $\sqrt{\sum_i \frac{(o_i-m_i)^2}{n}}$), bias ($\sum_i \frac{o_i-m_i}{n}$) and the coefficient of determination ($r^2 = 1 - \frac{\sum_i (o_i-m_i)^2}{\sum_i (o_i-\overline{o_i})^2}$) for the time series generated by the different tunings (o_i is observed and m_i is modelled).

3 Model tuning

The model was optimised for Hyytiälä with the AM algorithm using three different timescales: seasonal, daily and half-hourly tuning, which are described below. Tuning was done on a powerful laptop with an Intel Core i7-3520M processor. We removed unwanted output streams from the model and tweaked the I/O. With a single core the spin-up generation takes approximately 150 s, the run through calibration period with daily output takes 20 s and with half-hourly output 320 s. We used daily output also for the seasonal tuning.

3.1 Seasonal tuning

The fundamental motivation for the seasonal tuning is to ensure that the model reproduces the observed growing season maximum of LAI, since we have previously noticed that JSBACH underestimates LAI at the site level (even the default value of Δ_{max} is lower than the measured maximum for Hyytiälä). The reason for this approach was to enhance the vegetation transpiration and to emphasise the model response to precipitation. We also want to ensure that the model performs adequately well in terms of seasonal cumulative GPP and ET. The seasonal tuning was done in three consecutive steps each using the cost function (Eq. 1). The procedure is as follows:

1. All three class I parameters are tuned with four independent chains each consisting of 3000 samples. This step required a 30-year spin-up for each sample separately.
2. Class II and III parameters are each separately tested with 24 evenly separated values for an extensive range and those nine parameters that did not yield a negligible difference in the maximal and minimal values in the objective function are tuned. The consequent tuning was done with eight independent chains each consisting of 10 000 samples. A single spin-up, common for all samples, used optimal parameter values from step 1 and default values for the rest of the parameters.
3. All the previously tuned 12 parameters with eight independent chains each consisting of 10 000 samples are returned. Initial proposal covariance was generated from previous step and spin-up was generated separately for each sample.

At the end of seasonal tuning, class I parameters were fixed and a single spin-up was generated to be used with daily and half-hourly tuning. This approach is computationally justifiable (as we do not have to rerun the spin-up at each iteration of the algorithm) and is also acceptable from a modelling point of view since the robust site level scaling has already been done. The vegetative fraction of a grid cell remained at its default value of 0.52 and the carboxylation rate at 25 °C was lowered to 45.0 (and the electron transport rate to 85.5).

3.2 Daily and half-hourly tuning

The difference in daily and half-hourly tuning is the time interval used in the model output and observations in the cost function (Eq. 2). For both tuning runs we first tested the response of class II and III parameters against the cost function (Eq. 2) and removed those parameters that yielded only negligible or no response (as in step 2 in "Seasonal tuning"). The rest of the parameters (12) were then tuned using eight independent chains each consisting of 10 000 samples.

It should be noted that even though the cost function (Eq. 2) formulation is the same for daily and half-hourly tun-

Table 2. The highest correlations between parameters.

Tuning	Parameters		r
Seasonal	f_{C3}	w_{tsp}	0.49
	T_{alt}	α_q	0.40
Daily	f_{C3}	w_{tsp}	0.52
	w_{dr}	w_{tsp}	0.52
	T_{alt}	T_{ps}	−0.48
	T_{alt}	S_{min}	0.47
Half-hourly	f_{C3}	w_{tsp}	0.68
	p_{int}	w_{skin}	−0.44

ing, the values of the cost function are not directly comparable. Half-hourly tuning uses 48 values per day, and the resulting diurnal pattern resembles the form of the normal distribution. In daily tuning we use an average of these values. In practice, the component and cost function values will be higher for half-hourly tuning.

3.3 Tuning for Sodankylä

After tuning the model for Hyytiälä we took the parameter set from seasonal tuning and re-tuned only the maximum LAI parameter (Δ_{max}) with the cost function (Eq. 1) for Sodankylä. This was done because the measured LAI for Sodankylä is approximately half of that for Hyytiälä. The other parameter values were taken from the respective Hyytiälä tuning runs and spin-ups were generated similarly to Hyytiälä spin-ups so that we could use the Sodankylä results to validate the tuning process.

4 Results and discussion

The parameters and cost function components involved in the different phases of the optimisation process need to be studied before the performance of the optimisation method can be evaluated.

As noted above, we decided to reject the unreliable wintertime ET values. This masking leaves out the start of the growing season, which reduces the reliability of some of the tuned parameters, including all the LoGro phenology model parameters (class III), which mostly affect the timing of the spring event and regulate the development of the LAI towards the peak season. However, as a result of the tuning processes, all the analysed parameters were revealed to have unimodal posterior probability distributions, with different skewness and deviations.

We analysed the correlations and effectiveness of the parameters in the seasonal, daily and half-hourly optimisations on the Hyytiälä site for the calibration period. We also analysed the contributions from the cost function components referring to ET, GPP and LAI and generated the time series

Table 3. Significant components of principal component analysis for the different tunings. The given parameters are the most dominant within the component and the ratio is how many times larger the factor related to the first parameter is when compared to that of the second. Coverage reveals how much of the component is accounted for by the given parameters (sum of the weights of given vector components).

Component	Weight	Parameters		Ratio	Coverage
Seasonal 1	0.996	w_{hum}	w_{skin}	2.1	> 99 %
Daily 1	0.717	T_{ps}	w_{skin}	1.4	> 99 %
Daily 2	0.261	w_{hum}	w_{tsp}	2.3	> 99 %
Half-hourly 1	0.530	T_{ps}	–	–	> 99 %
Half-hourly 2	0.310	w_{skin}	w_{hum}	1.7	96 %
Half-hourly 3	0.121	T_{alt}	–	–	> 99 %

and daily cycles of GPP and ET for both Hyytiälä and Sodankylä sites. For all these examinations, individual spin-ups were generated using the optimised parameter values.

The parameter correlations (Table 2) do not reveal much information, which is common for larger systems where the underlying parameter dependencies are more complex. Usually more sophisticated methods are used to analyse the parameters, but we omit these examinations here since pairwise Kernel density estimates (Fig. 1) did not reveal any new insights.

The strongest correlation was between the ratio of leaf internal CO_2 concentration to external CO_2 (f_{C3}) and fraction of soil moisture above which transpiration is unaffected by soil moisture stress (w_{tsp}) in all the tunings. This positive correlation strengthens as we increase the temporal resolution (and the complexity of the underlying cost function). This is due to the carbon assimilation that is limited not only by the amount of carbon available but also by a linear water stress factor (which takes the value of zero at the wilting point (w_{wilt}) and one at the w_{tsp}), which is checked at each time step. Most of the other parameters with high correlations are those of the LoGro phenology model, where we would expect high correlation since the parameters are intimately connected.

Approximately half of the parameters with high correlation are also the least identifiable (Table 3) with the given data and cost function. This means that the values these parameters acquire, as a result of the tuning process, are the most unreliable – it does not reflect on the parameters contribution to the cost function. The PCA merely highlights where most of the parametric unreliability lies.

The PCA analysis revealed that most of the unreliability is explained by a handful of parameters. Disregarding those of the LoGro phenology model, the two most dominantly unreliable parameters in every tuning were the fraction depicting relative humidity based on soil dryness (w_{hum}) and the maximum field capacity of the skin reservoir (w_{skin}). Both of these parameters affect the amount of water available for

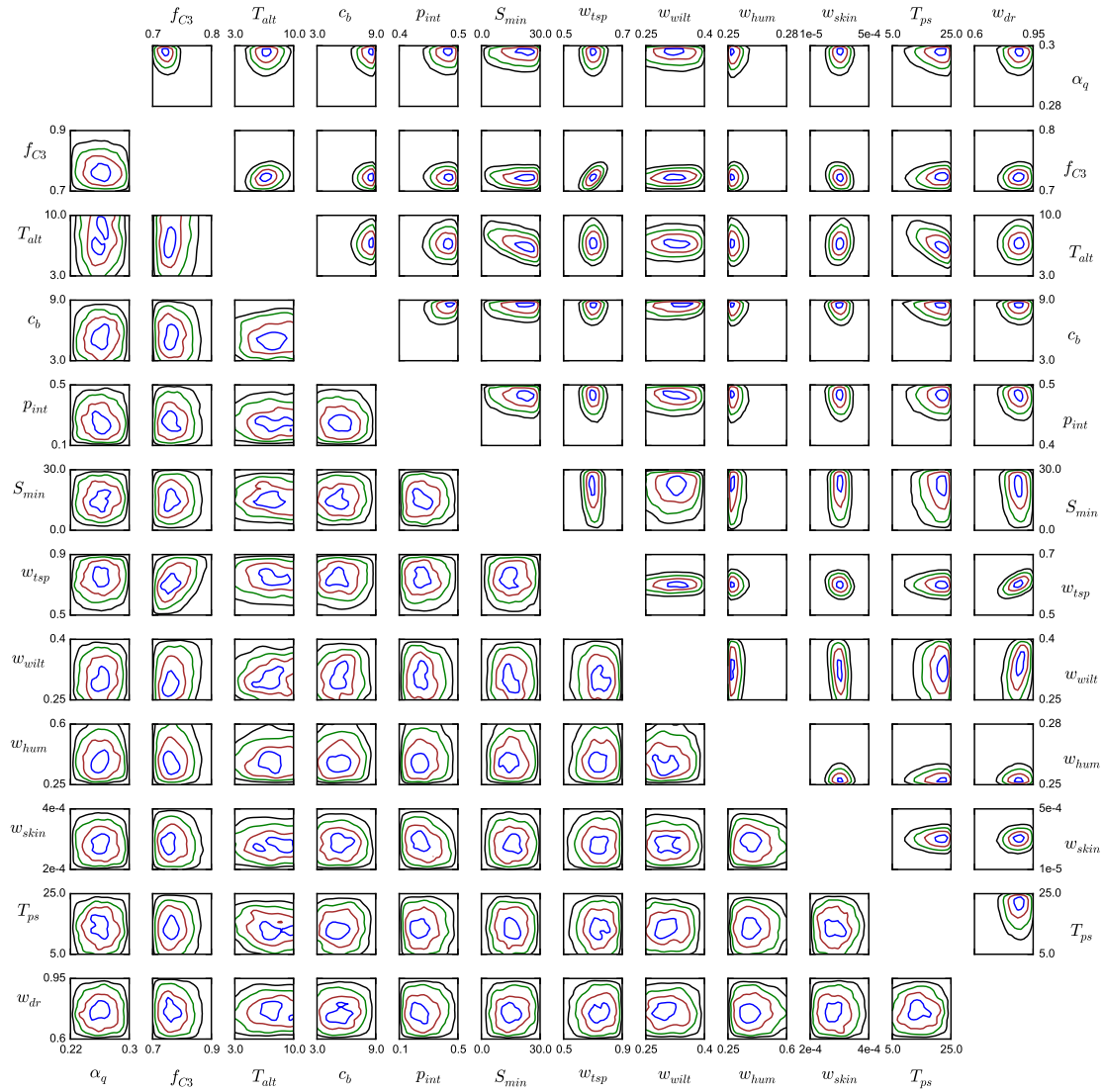

Figure 1. Kernel density estimates of the last 20 000 parameter samples with daily (upper triangle) and half-hourly tunings. The contours correspond to densities in a two-dimensional Gaussian distribution (μ_x, $\mu_y = 0$, σ_x, $\sigma_y = 1$) with 2σ (black), 1.5σ (green), σ (brown), 0.5σ (blue).

evaporation from bare soil and are both subject to changes in other parameters. Bare soil evaporation is also dominated by vegetative transpiration, which explains why these two parameters are the most unreliable.

4.1 The parameters and their relative effectiveness

The default and optimised parameter values from the different tuning metrics are presented in Table 4 along with their relative effectiveness. The reference values for seasonal tuning are the default values. Since we fixed class I parameters with seasonal tuning, the realistic reference values for daily and half-hourly tunings are the seasonal parameter val-

Table 4. Default and optimised parameter values using the last 20 000 samples (if no value is given, the parameter was not part of that tuning, and the default value was used instead). The percentage next to a parameter value is the effectiveness of that parameter for that tuning. The reference values for seasonal tuning are the default values and for daily and half-hourly tunings the seasonal values.

Parameter		Default	Seasonal		Daily		Half-hourly	
α_q		0.28	0.26	7 %	0.30	3 %	0.27	1 %
c_b		5.0	–	–	8.8	7 %	5.0	0 %
f_{C3}		0.87	0.88	8 %	0.72	70 %	0.76	68 %
p_{int}		0.25	0.27	1 %	0.49	4 %	0.27	0 %
w_{dr}		0.9	0.79	14 %	0.87	1 %	0.75	−1 %
w_{hum}		0.5	0.54	1 %	0.25	14 %	0.37	22 %
w_{pwp}		0.35	0.28	10 %	0.34	0 %	0.31	−1 %
w_{skin}	[m]	2.0×10^{-4}	3.1×10^{-4}	6 %	3.0×10^{-4}	0 %	2.2×10^{-4}	6 %
w_{tsp}		0.75	0.64	53 %	0.60	1 %	0.75	3 %
T_{alt}	[°C]	4.0	8.1	0 %	6.9	1 %	6.9	2 %
S_{min}	[°C]	10.0	–	–	23.0	−0 %	14.7	−0 %
T_{ps}	[°C]	10.0	–	–	21.0	−0 %	12.4	−0 %

ues. Here we note that using one spin-up for all daily and half-hourly optimisation runs is computationally justifiable but generates errors as the general spin-up differs from those generated by the optimised parameters. These errors are relatively small but give rise to, e.g., the negative relative effectiveness values in daily and half-hourly parametrisations.

Most seasonally tuned parameters are near their default values and the most effective parameters are the fraction of soil moisture above which transpiration is unaffected by soil moisture stress (w_{tsp}), the fraction of soil moisture at permanent wilting point (w_{pwp}) and the fraction of field capacity above which fast drainage occurs (w_{dr}). For daily and half-hourly tunings the most important parameters are the ratio of leaf internal CO_2 concentration to external CO_2 (f_{C3}) and the fraction depicting relative humidity (w_{hum}). It should be noted that w_{hum} was one of the least identifiable parameters for seasonal tuning. Taking into account the importance of these parameters on transpiration and soil moisture estimations, we took a closer look at modelled soil moisture and evapotranspiration components for the calibration period (taking into account only values from May to September for each year as explained at the end of Sect. 2.5.

When we compare the model output streams with seasonal against those with default parametrisation, we notice that the average evapotranspiration for the calibration period has increased 15 %. Most of this is due to not only added transpiration (18 % increase) but also increased evaporation (6 %). In addition drainage was accelerated by 11 %. These increases are mostly compensated by a 15 % reduction in average soil moisture. In addition soil moisture values that are under the limit when transpiration is affected by soil moisture stress (below the value of w_{tsp}) increased 2.3 %.

The daily and half-hourly tunings lower the average evapotranspiration by 22 and 35 % respectively, when compared to the seasonal values. Transpiration is decreased by 28 and 37 %, whereas evaporation is increased by 0.5 % and decreased by 28 % respectively, for daily tuning and half-hourly tuning. Soil moisture is increased by 11 and 8 % and the amount of values below w_{tsp} is decreased by 62 % for daily tuning and increased by 7 % for half-hourly tuning. Out of curiosity, both the adjustment parameter in stability functions (c_b) and the fraction of precipitation intercepted by canopy (p_{int}) have been significantly increased with daily tuning and returned to seasonally tuned values with half-hourly tuning.

4.2 The cost function components

Using the optimised values (parametrisations), we calculated the components of each cost function for Hyytiälä calibration period and Hyytiälä and Sodankylä validation period (Table 5).

First, we note that with the default parameters L_1 dominates cf_1 for Hyytiälä and contributes approximately 90 % to its value. As expected the L_1 for Sodankylä is not as dominant as for Hyytiälä since the measured maximum of LAI for Hyytiälä is roughly half as large as for Sodankylä, which directly lowers the LAI component in cost function (Eq. 1). The L_1 contribution is significantly reduced with the seasonally tuned parameters as was our intention and even though LAI plays no part in daily and half-hourly tunings, the differences in the maximum value are negligible.

Second, the value of the E_1 component (error in seasonal ET) with default parametrisation is significantly increased in daily and especially half-hourly parametrisations. Simultaneously the value of G_1 is significantly lowered. The component values for seasonal parametrisation are better than the default values with the exception of E_1 for Hyytiälä validation period.

Third, for the cost function (Eq. 2) the pairwise ratio of dominating E_i or G_i components in all tunings is 5 : 1. On average E_2/E_3 contributes to approximately 60% of cf_2/cf_3.

Table 5. Cost function components for each parametrisation for Hyytiälä calibration (HC), Hyytiälä validation (HV) and Sodankylä validation (SV) periods. L_1, E_1 and G_1 are the LAI, ET and GPP components in cost function (Eq. 1), represented by cf_1 and used for seasonal tuning. Likewise E_2 and G_2 are the components in cost function (Eq. 2) for daily values (cf_2), whereas E_3 and G_3 are for half-hourly values (cf_3). Note that the values of cf_2 and cf_3 are not directly comparable.

		L_1	E_1	G_1	E_2	G_2	E_3	G_3	cf_1	cf_2	cf_3
HC	Default	0.396	0.021	0.036	0.306	0.191	1.126	0.681	0.45	0.50	1.8
	Seasonal	5×10^{-5}	1.7×10^{-4}	5.7×10^{-6}	0.343	0.161	1.326	0.720	2.3×10^{-4}	0.50	2.0
	Daily	7.4×10^{-5}	0.055	1.4×10^{-4}	0.206	0.149	0.906	0.683	0.06	0.36	1.6
	Half-hourly	1.0×10^{-4}	0.128	5.4×10^{-3}	0.276	0.151	0.864	0.661	0.13	0.43	1.5
HV	Default	0.396	0.002	0.028	0.226	0.157	1.027	0.479	0.43	0.38	1.5
	Seasonal	9.3×10^{-5}	0.011	7.5×10^{-4}	0.300	0.134	1.370	0.459	0.01	0.43	1.8
	Daily	1.4×10^{-4}	0.007	3.5×10^{-4}	0.164	0.124	0.981	0.446	7×10^{-3}	0.29	1.4
	Half-hourly	1.1×10^{-4}	0.058	2.9×10^{-3}	0.182	0.118	0.748	0.412	0.06	0.30	1.2
SV	Default	0.108	4.0×10^{-3}	0.140	0.423	0.596	1.660	1.795	0.25	1.02	3.5
	Seasonal	5.9×10^{-3}	1.8×10^{-5}	0.068	0.467	0.411	1.786	1.429	0.07	0.88	3.2
	Daily	6.1×10^{-3}	0.063	0.048	0.289	0.352	1.258	1.294	0.12	0.64	2.6
	Half-hourly	5.9×10^{-3}	0.164	0.022	0.379	0.290	1.246	1.185	0.19	0.67	2.4

This translates to ET being twice as significant as GPP in the cost function (Eq. 2). The main reason for ET dominating GPP is that ET is more erratic in comparison to GPP and the residuals of ET (divided by the mean value) are larger than the residuals of GPP. The daily and half-hourly tunings themselves work as intended as they lower the corresponding cost function value. It is noteworthy to mention that the G_2 component gets its lowest value for both validation periods with the half-hourly parametrisation even though G_2 calculates GPP errors on a daily scale.

Lastly, we examine how the algorithm and cost functions have performed. The best parameter set (the lowest cost function value) for a given cost function, in each of the three different periods (HC, HV, SV), is the same as that used in the corresponding tuning process. For example the lowest value for cf_1 (the cost function for seasonal tuning) in Sodankylä validation period (0.07) coincides with the seasonally tuned parameters. This is expected as the tuning process aims to be the "best" parameter value, which reassures us that no gross mistakes (human errors) have been made. The relation holds true for every cost function with the exception of cf_1 for Hyytiälä validation period, where the lowest value is reached with the daily tuned parameters (we note that the absolute difference between daily and seasonally tuned parameters is small). Hence we can confidently state that the algorithm and cost functions have performed as intended, especially since the optimised parameters work for Sodankylä as well, where no optimisation (besides the site-specific maximum of LAI) was applied.

4.3 Time series

The overall structure of the model time series was not affected by the parametrisations obtained with different tunings (Figs. 2 and 3). Some time series characteristics have been enhanced and others reduced but the timing of the peaks and dips in GPP and ET are the same as before. The corresponding RMSE and bias estimates are given in Table 6. In comparison we estimated the PRELES model biases for Hyytiälä from Fig. 5 in Peltoniemi et al. (2015b). These estimates give a bias of $0.81 \times 10^{-6}\,\mathrm{kg\,m^{-2}\,s^{-1}}$ ($0.07\,\mathrm{mm\,m^{-2}\,day^{-1}}$) for ET and $-1.45 \times 10^{-7}\,\mathrm{mol\,[CO_2]\,m^{-2}\,s^{-1}}$ ($-0.15\,\mathrm{g(C)\,m^{-2}\,day^{-1}}$) for GPP. Additionally, the coefficient of determination (r^2) for GPP in Hyytiälä is in the range of 0.74–0.76 for all tunings, whereas the values reported in literature range from 0.68 (Trusilova et al., 2004) to 0.96 (Peltoniemi et al., 2015b) with most above 0.9 (Aalto et al., 2004; Dursma et al., 2009). For additional comparisons see also Abramowitz et al. (2007). Note that our estimates are calculated using only values from the beginning of May to the end of September.

The best seasonal performance was obtained by seasonal tuning as we previously noted from the cost function components (Table 5). Even though the optimisation is done on the seasonal level, especially the GPP cycle is noticeably improved from that generated by the default parameters. This tuning also gives rise to the most stable (least fluctuating) water use efficiency (WUE), when calculated as a pointwise ratio of GPP and ET. We use WUE here only as a diagnostic variable to examine the balance between the GPP and ET.

When compared to the seasonal tuning, the daily tuning is worse on the seasonal scale and lowers both the ET and GPP cycles. WUE follows the observations better but starts to give rise to some fluctuation. With half-hourly tuning, this behaviour is further enhanced and especially ET is lowered to too low levels, which manifests the high WUE values. The

Table 6. RMSE and bias of ET and GPP calculated from half-hourly data for first two summers of the validation period for Hyytiälä (corresponding to Fig. 2) and last two summers of the validation period for Sodankylä (corresponding to Fig. 3).

	ET ($kg\,m^{-2}\,s^{-1}$)				GPP ($mol\,(CO_2)\,m^{-2}\,s^{-1}$)			
	Hyytiälä		Sodankylä		Hyytiälä		Sodankylä	
	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias
Default	2.03×10^{-5}	-1.31×10^{-6}	$2.27E \times 10^{5}$	2.31×10^{-6}	3.09×10^{-6}	8.77×10^{-7}	3.16×10^{-6}	-9.19×10^{-7}
Seasonal	2.37×10^{-5}	-4.32×10^{-6}	2.35×10^{-5}	1.09×10^{-6}	3.10×10^{-6}	-2.00×10^{-7}	2.89×10^{-6}	-5.97×10^{-7}
Daily	2.03×10^{-5}	-0.74×10^{-6}	2.06×10^{-5}	5.00×10^{-6}	3.06×10^{-6}	-1.07×10^{-7}	2.74×10^{-6}	-4.57×10^{-7}
Half-hourly	1.69×10^{-5}	2.77×10^{-6}	2.04×10^{-5}	7.14×10^{-6}	2.94×10^{-6}	3.39×10^{-7}	2.67×10^{-6}	-2.79×10^{-7}

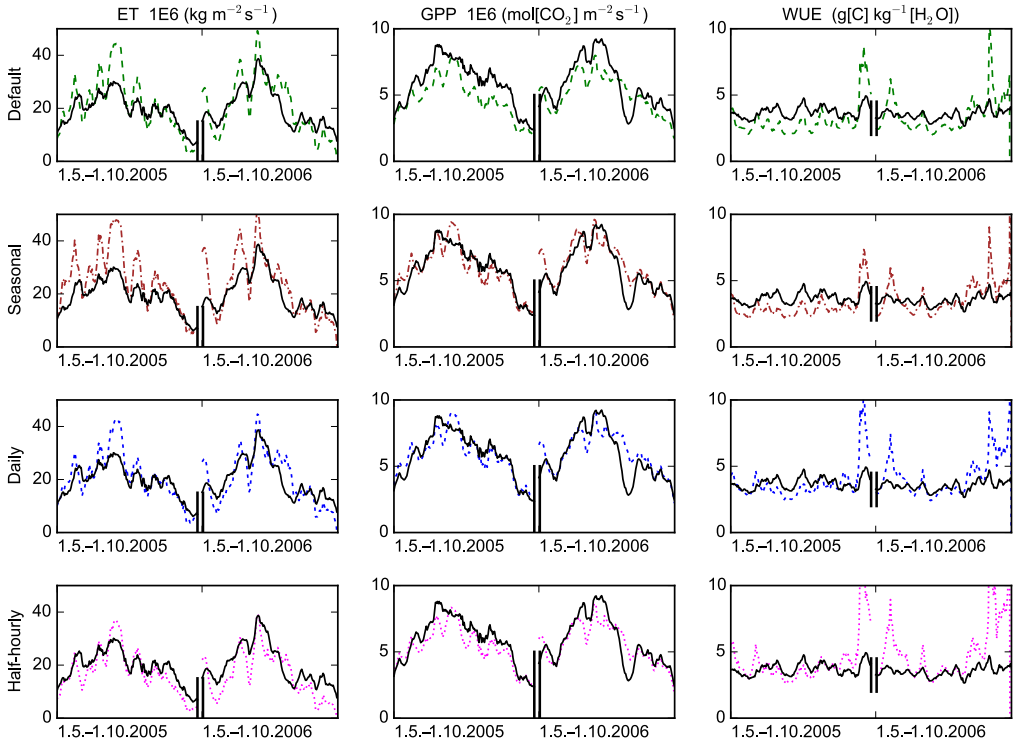

Figure 2. Hyytiälä 7-day-running mean time series for different tunings for the first two summers of the validation period. Solid black line represents the observations.

worsening in the model time series with daily and half-hourly tunings are explained by biases in the diurnal cycle.

4.4 Diurnal cycles

Average diurnal cycles with different parametrisations (Fig. 4) show that modelled night-time ET values are too low when compared to the observed and this behaviour was not affected by the tunings. Low night-time values are compensated by too high midday values in the default and seasonal tuning so that the average daily and seasonal values are on an acceptable level. For the daily and half-hourly tuning, the algorithm lowers the daytime values, which results in too low average daily and half-hourly values. It is noteworthy to mention that with the default setting we get too low GPP for Hyytiälä but too high GPP for Sodankylä. The unrealistic wintertime and the biased night-time ET values actually have the same origin. Since we do not have the coupling from the land surface model (LSM) back to the atmosphere, we get an erroneous energy balance as we lose the energy released by condensation.

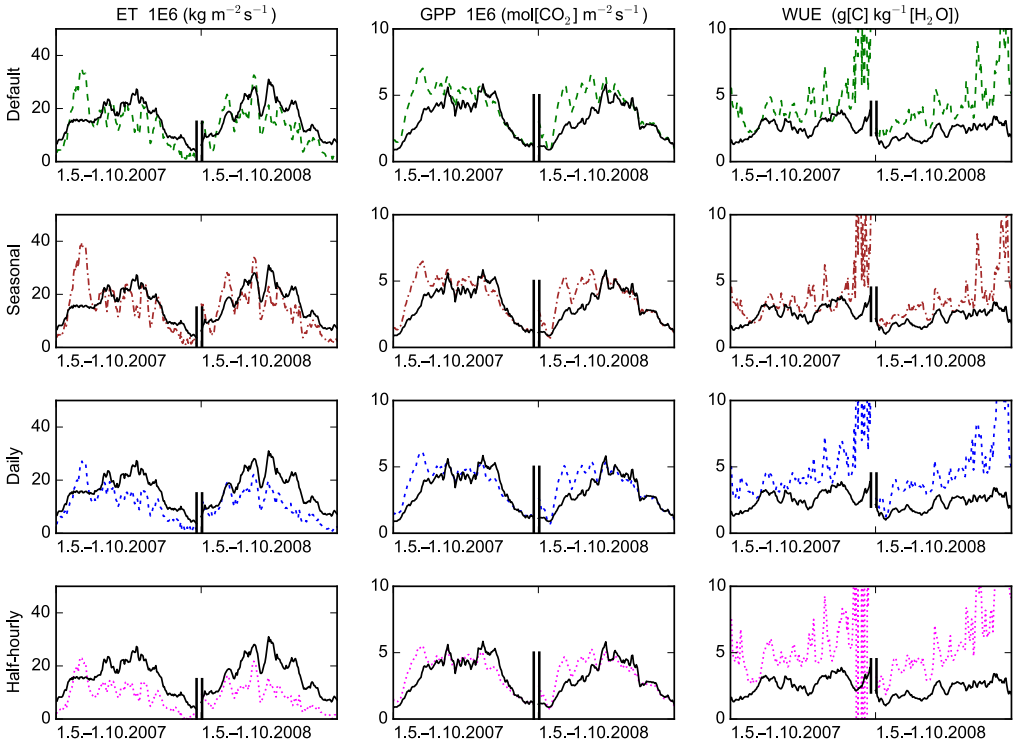

Figure 3. Sodankylä 7-day-running mean time series for different tunings for the last two summers of the validation period. Solid black line represents the observations.

Disregarding the default parametrisation we notice that seasonal parametrisation show the highest values, daily in the middle and half-hourly show the lowest values. Daily parametrisation reproduces the observations for average diurnal cycle better than the others in every occasion except the GPP for Sodankylä, where half-hourly tuning is better (verified by pointwise RMSE from the average diurnal cycle). We also notice that Sodankylä daily patterns, and to some extent Hyytiälä as well, are slightly out of phase. Our current understanding is that this is (at least partly) due to a slightly misaligned sensor (which can cause significant errors on high latitudes), measuring radiation fluxes. Fortunately this affects mainly the cost function for half-hourly tuning since it is the only one operating on the densest half-hourly timescale.

4.5 Dry event

Dry period in the summer 2006 can be clearly located by the massive drawdown in observed GPP, and to a lesser extent in ET, at Hyytiälä (Fig. 2). In a closer look at this event (Fig. 5) it is evident that none of our parametrisation schemes were able to capture it correctly. As it was with the time series, the overall structure of the daily time series during this event remains the same (there are no divergent aspects in the model output between the different tunings).

During the drought event (defined here as 31 July–15 August 2006), the soil moisture is on average 27 % lower for default, daily and half-hourly tuning and 40 % lower for seasonal tuning when compared to the corresponding values from other years – seasonal tuning has the lowest overall soil moisture. During this event the modelled soil moisture decreases monotonically for all tunings and reaches the lowest values on 13 August, after which it starts to rise. During the period the modelled ET and GPP are predominantly higher than the observations. WUE on the other hand follows the "observations" remarkably well and deviates from the observed only towards the end of the event when modelled ET drops to near-zero values, coinciding with the lowest modelled soil moisture values. Gao et al. (2016) examined deviation in the dependencies of GPP and ET to vapour pressure deficit (VPD) between model and observation results under the most severe soil moisture stress conditions at the end of the prolonged period of soil water scarcity (that occurred in 2006). This can be attributed to the lack of ex-

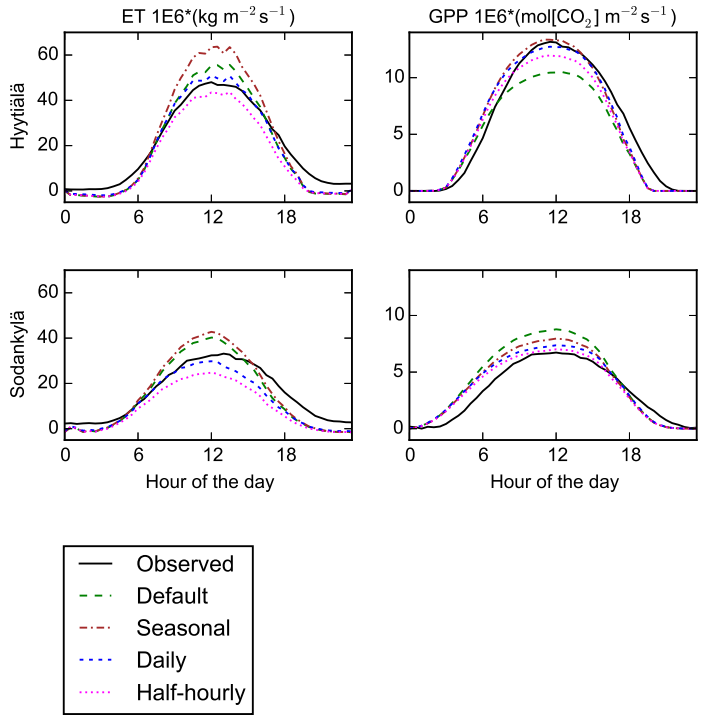

Figure 4. Average diurnal cycle from May to September for the validation period.

plicit dependence of the modelled stomatal conductance on the atmospheric humidity.

5 Conclusions

Initially we tuned the model to produce near-measured seasonal ET, GPP and especially maximum LAI to enhance the vegetation transpiration and to emphasise the response to precipitation. This was done successfully with seasonal tuning in the hopes of bringing forth the underlying model responses to dryness. With the consecutive daily and half-hourly tunings, we managed to improve the average diurnal cycles of both ET and GPP, but failed in reproducing the low ET and GPP levels during the dry event in 2006. Effectively we first (seasonal tuning) transferred water from soil moisture into (too high levels of) ET, and later (with daily and half-hourly tunings) transferred some of it back.

In addition to the maximum LAI ($\Delta_{\max}$) and maximum carboxylation rate ($V_{\mathrm{C,max}}$), the most effective parameters in the seasonal tuning were the fraction of soil moisture above which transpiration is not affected by soil moisture stress (w_{tsp}) and the critical fraction of field capacity above which fast drainage occurs for soil water content (w_{dr}). The reduction in ET and GPP was mostly accounted for by lowering the approximate ratio of leaf internal CO_2 concentration to external CO_2 (f_{C3}), which reduces the amount of carbon available for photosynthesis. For daily tuning ET was further reduced by the increase of the fraction of precipitation intercepted by canopy (p_{int}) and lower relative humidity fraction (w_{hum} – air humidity is based on soil dryness).

Despite the fact that we were unable to enhance the dry response of the model, we are confident in saying that the algorithm itself worked well and performed as intended with the daily tuning providing the most reduction in model–data mismatch. We optimised 12 parameters simultaneously (with daily and half-hourly tunings) using eight fairly short chains (8000 samples). With daily tuning the resulting estimates are well matured, but with half-hourly tuning the parameter deviations are larger (which is probably due modelling inefficiencies and noise in measurements). Nevertheless, all optimisation procedures worked well with regard to what was optimised (seasonality, daily averages or diurnal cycle).

Recently, Knauer et al. (2015) found canopy conductance formulation to be a key factor in prescribing the transfer of carbon and water between terrestrial biosphere and the lower atmosphere. Additionally, Gao et al. (2016) found that during a prolonged period of soil water scarcity, the lack of explicit

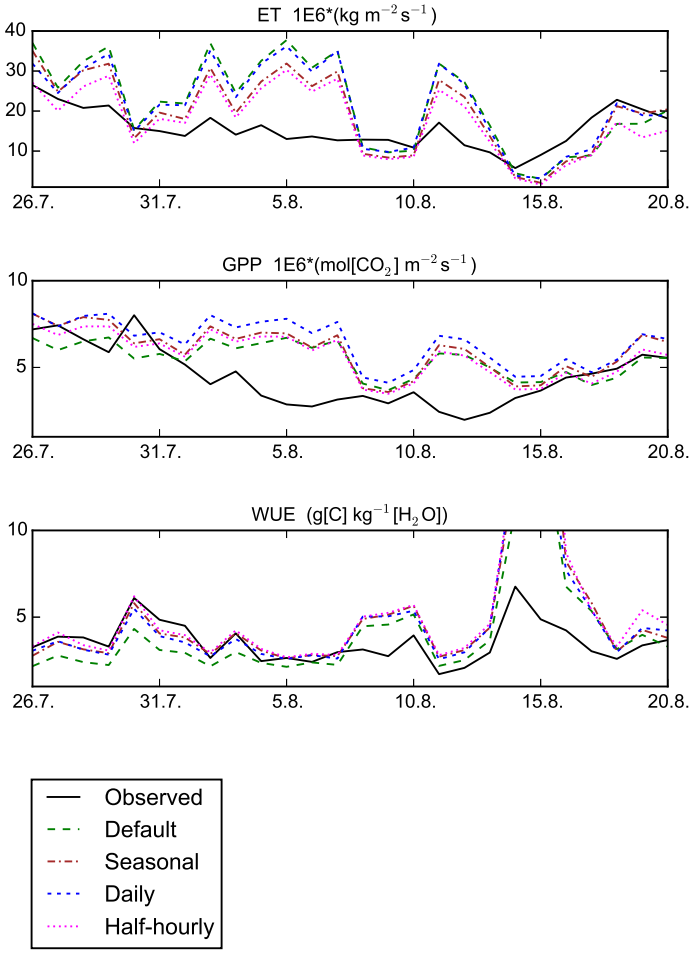

Figure 5. Daily averages for ET, GPP and WUE on a dry event in 2006 for Hyytiälä.

dependence of the stomatal conductance on the atmospheric humidity is one of the contributing factors to this issue. Further studies into enhancing the dry response in JSBACH are needed and these studies should reflect these latest findings.

6 Data availability

The measurement data required to run and tune the model can be procured from the FLUXNET database (doi:10.17616/R36K9X). The JSBACH model is available to the scientific community under a version of the Max Planck Institute for Meteorology Software License Agreement (http://www.mpimet.mpg.de/en/science/models/license/). For any questions regarding the simulations data, we encourage you to contact the author at jarmo.makela@fmi.fi.

Appendix A: Parametric equations within JSBACH

In this Appendix we present the main equations that the parameters in this study affect.

A1 Logistic Growth Phenology model

The parameters from the LoGro-P model, which are of interest here, are mainly used to determine the spring event for JSBACH. The maximum all-sided leaf area index ($\Delta_{\max}$) is also part of this model; hence, we introduce this first and then deal with the spring event. $\Delta_{\max}$ is used to calculate LAI at each time step by a logistic equation (Eq. A1). Here k is the growth and p the shedding rate, both of which further depend on temperature and soil moisture.

$$\frac{\mathrm{d}\Delta}{\mathrm{d}t} = k\Delta\left(1 - \frac{\Delta}{\Delta_{\max}}\right) - p\Delta \tag{A1}$$

To determine the date of the spring event, we first introduce a few additional variables, namely, the heat sum ($S_T(d)$), the number of chill days ($C(d)$) and the critical heat sum ($S_{\mathrm{crit}}(d)$). Also $T(d)$ denotes the mean temperature at day d.

$$S_T(d) = \sum_{d'=d_0}^{d} \max\left(T(d') - T_{\mathrm{alt}}, 0\right) \tag{A2}$$

Heat sum $S_T(d)$ cumulates the amount of "heat" above the parameter T_{alt} after the previous growing season. The actual starting date d_0 of the summation need not be known since it is enough to start the summation "reasonably late" after the last growth season.

$$C(d) = \sum_{d'=d_{\mathrm{a}}}^{d} H\left(T_{\mathrm{alt}} - T(d)\right) \tag{A3}$$

The number of chill days is calculated as the number of days when the mean temperature is below T_{alt}. Here $H()$ denotes the Heaviside step function and the summation starts at the day (d_{a}) of the last autumn event.

$$S_{\mathrm{crit}}(d) = S_{\min} + S_{\mathrm{range}} e^{-C(d)/C_{\mathrm{decay}}} \tag{A4}$$

The critical heat sum (S_{crit}) decreases as the number of chill days $C(d)$ increases. The spring event happens when

$$S_T(d) \geq S_{\mathrm{crit}}(d). \tag{A5}$$

Pseudo soil temperature ($T_{\mathrm{s}}(t)$) at time t is calculated as an average air temperature (T) with an exponential memory loss (T_{ps}). Pseudo soil temperature is used in determining the autumn event (when it falls below a certain threshold). In the equation N is the normalisation constant and τ is the length of a time step.

$$T_{\mathrm{s}}(t) = \frac{1}{N} \sum_{n=-\infty}^{t} T(n) e^{-(t-n)\frac{\tau}{T_{\mathrm{ps}}}} \tag{A6}$$

A2 Photosynthesis

The Farquhar model is based on the observation that the assimilation rate in the chloropast is limited either by the carboxylation rate (V_{C}) or the transport rate (J_{E}) of two electrons freed during the photoreaction. The total rate of carbon fixation A is given by the following equation, where R_{d} is the dark respiration:

$$A = \min\left(V_{\mathrm{C}}, J_{\mathrm{E}}\right) - R_{\mathrm{d}}. \tag{A7}$$

Oxygenation of the Rubisco molecule reduces the carboxylation rate, which is given as

$$V_{\mathrm{C}} = V_{\mathrm{C,max}} \frac{C_{\mathrm{i}} - \Gamma_*}{C_{\mathrm{i}} + K_{\mathrm{C}}\left(1 + O_{\mathrm{i}}/K_{\mathrm{O}}\right)}. \tag{A8}$$

Here C_{i} and O_{i} are the leaf internal CO_2 and O_2 concentrations, $\Gamma_\star$ is the CO_2 compensation point, K_{C} and K_{O} are Michaelis–Menten constants parametrising the dependence on CO_2 and O_2 concentrations. Furthermore, leaf internal CO_2 concentration depends on the external concentration C_{E} by

$$C_{\mathrm{i}} = f_{\mathrm{C3}} C_{\mathrm{E}}. \tag{A9}$$

Likewise the electron transport rate is given as

$$J_{\mathrm{E}} = J(I) \frac{C_{\mathrm{i}} - \Gamma_*}{4\left(C_{\mathrm{i}} + 2\Gamma_*\right)}. \tag{A10}$$

Here $J(I)$ is a function of radiation intensity I in the photosynthetically active band, the maximum electron transport rate $J_{\max}$ and the quantum efficiency for photon capture α_{q}.

$$J(I) = J_{\max} \frac{\alpha_{\mathrm{q}} I}{\sqrt{J_{\max}^2 + \alpha_{\mathrm{q}}^2 I^2}} \tag{A11}$$

A3 Soil water

In JSBACH, the soil water budget is based on several reservoirs (skin, soil, bare soil, rain intercepted by canopy, etc.) and the different formulations are plentiful. We present here only the most crucial of these. Changes in soil water (w_{s}) due to rainfall (R), evapotranspiration (ET), snowmelt (M), surface runoff (R_{s}) and drainage (D) are calculated with a geographically varying maximum field capacity (w_{fc}).

$$\rho \frac{\partial w_{\mathrm{s}}}{\partial t} = \left(1 - p_{\mathrm{int}}\right) R + \mathrm{ET} + M - R_{\mathrm{s}} - D \tag{A12}$$

The interception parameter (p_{int}) also affects the amount of water intercepted by vegetation and bare soil that further affects evaporation, etc. The skin reservoir is limited by w_{skin} and excess water is transferred to soil water. Likewise when the soil water content (in relation to maximum field capacity) is greater than parameter w_{dr}, the excess water is rapidly

drained (in addition to the limited drainage below this threshold).

Evaporation from wet surfaces (E_{ws}) depends on air density (ρ), specific humidity (q_a), saturation-specific humidity (q_s) at surface temperature (T_s) and pressure (p_s) and aerodynamic resistance ($r_a = C_h|v_h|^{-1}$; these are heat transfer coefficient and horizontal velocity).

$$E_{ws} = \rho \frac{q_a - q_s(T_s, p_s)}{r_a} \tag{A13}$$

Transpiration from vegetation (T_v) is likewise formulated but additionally depends on the stomatal resistance of canopy (r).

$$T_v = \rho \frac{q_a - q_s(T_s, p_s)}{r_a + r} \tag{A14}$$

The stomatal resistance is given as a minimal stomatal resistance of the canopy without water stress (r_{min}, depends on photosynthetically active radiation and LAI) divided by a water stress factor (f_{ws}). That is $r = r_{min}/f_{ws}$. The water stress factor depends on how much water is in the soil in relation to the maximum field capacity ($w_f = w_s/w_{fc}$) when compared to the limit when transpiration is no longer affected by soil moisture stress (w_{tsp}) and the permanent wilting point (w_{pwp}).

$$f_{ws} = \begin{cases} 1 & w_f \geq w_{tsp} \\ \dfrac{w_f - w_{pwp}}{w_{tsp} - w_{pwp}} & w_{pwp} \leq w_f \leq w_{tsp} \\ 0 & w_f \leq w_{pwp} \end{cases} \tag{A15}$$

Evaporation from dry bare soil (E_s) is similarly defined as

$$E_s = \rho \frac{q_a - h q_s(T_s, p_s)}{r_a} \tag{A16}$$

Here h is relative humidity at the surface relative to soil dryness:

$$h = \max\left[w_{hum}\left(1 - \cos\left(\pi w_f\right)\right), \min\left(1, \frac{q_a}{q_s(T_s, p_s)}\right)\right]. \tag{A17}$$

The total evapotranspiration is a weighted average of E_{ws}, T_v and E_s, where the weights are based on, e.g., fill levels of reservoirs (similar to w_f above) and vegetative fraction of the grid cell (veg_{max}).

Author contributions. Tuula Aalto, Heikki Järvinen, Tiina Markkanen and Stefan Hagemann chose the parameters in the optimisation process and provided support throughout the experiments. Mika Aurela and Ivan Mammarella provided knowledge on the observations. Jouni Susiluoto provided the algorithm test bed and Jarmo Mäkelä integrated the model into the test bed, ran the experiments and prepared the manuscript with contributions from all co-authors.

Acknowledgements. This work was funded by the European Commission's 7th Framework Programme, under grant agreement no. 282672, EMBRACE project, and the Nordic Centre of Excellence "Tools for Investigating Climate Change at High Northern Latitudes" (eSTICC) under the Nordic Top-Level Research Initiative. This work was also supported by the Academy of Finland Center of Excellence (no. 272041), ICOS-Finland (no. 281255) and ICOS-ERIC (no. 281250) funded by Academy of Finland. This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia and USCCC. The FLUXNET eddy covariance data processing and harmonisation was carried out by the ICOS Ecosystem Thematic Center, AmeriFlux Management Project and Fluxdata project of FLUXNET, with the support of CDIAC, and the OzFlux, ChinaFlux and AsiaFlux offices.

Edited by: O. Talagrand
Reviewed by: two anonymous referees

References

Aalto, T., Ciais, P., Chevillard, A., and Moulin, C.: Optimal determination of the parameters controlling biospheric CO_2 fluxes over Europe using eddy covariance fluxes and satellite NDVI measurements, Tellus B, 56, 93–104, doi:10.3402/tellusb.v56i2.16413, 2004.

Abramowitz, G., Pitman, A., Gupta, H., Kowalczyk, E., and Wang, Y.: Systematic Bias in Land Surface Models, J. Hydrol., 8, 989–1001, doi:10.1175/JHM628.1, 2007.

Aurela, M.: Carbon dioxide exchange in subarctic ecosystems measured by a micrometeorological technique, Finnish Meteorol. Inst. Contr., 51, 1–39, 2005.

Aurela, M., Lohila, A., Tuovinen, J., Hatakka, J., Riutta, T., and Laurila, T.: Carbon dioxide exchange on a northern boreal fen, Boreal Environ. Res., 14, 699–710, 2009.

Baldocchi, D.: Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future, Global Change Biol., 9, 479–492, doi:10.1046/j.1365-2486.2003.00629.x, 2003.

Boé, J. and Terray, L.: Uncertainties in summer evapotranspiration changes over Europe and implications for regional climate change, Geophys. Res. Lett., 35, L05702, doi:10.1029/2007GL032417, 2008.

Craiu, R., Rosenthal, J., and Yang, C.: Learn From Thy Neighbor: Parallel-Chain and Regional Adaptive MCMC, J. Am. Stat. Assoc., 104, 1454–1466, doi:10.1198/jasa.2009.tm08393, 2009.

Dalmonech, D., Zaehle, S., Schürmann, G., Brovkin, V., Reick, C., and Schnur, R.: Separation of the Effects of Land and Climate Model Errors on Simulated Contemporary Land Carbon Cycle Trends in the MPI Earth System Model version 1, J. Climate, 28, 272–291, doi:10.1175/JCLI-D-13-00593.1, 2015.

Duursma, R., Kolari, P., Perämäki, M., Pulkkinen, M., Mäkelä, A., Nikinmaa, E., Hari, P., Aurela, M., Berbigier, P., Bernhofer, C., Grünwald, T., Loustau, D., Mölder, M., Verbeeck, H., and Vesala, T.: Contributions of climate, leaf area index and leaf physiology to variation in gross primary production of six coniferous forests across Europe: a model-based analysis, Tree Physiol., 29, 621–639, doi:10.1093/treephys/tpp010, 2009.

Farquhar, G., von Caemmerer, S., and Berry, J.: A Biochemical Model of Photosynthetic CO_2 Assimilation in Leaves of C3 species, Planta, 149, 78–90, doi:10.1007/BF00386231, 1980.

Gao, Y., Markkanen, T., Aurela, M., Mammarella, I., Thum, T., Tsuruta, A., Yang, H., and Aalto, T.: Response of water use efficiency to summer drought in boreal Scots pine forests in Finland, Biogeosciences Discuss., doi:10.5194/bg-2016-198, in review, 2016.

Haario, H., Saksman, E., and Tamminen, J.: An adaptive Metropolis algorithm, Bernoulli, 7, 223–242, 2001.

Hagemann, S.: An improved land surface parameter dataset for global and regional climate models, Max Planck Institute for Meteorology Report 336, 1–28, https://www.mpimet.mpg.de/fileadmin/publikationen/Reports/max_scirep_336.pdf (last access: 7 December 2016), 2002.

Hagemann, S. and Stacke, T.: Impact of the soil hydrology scheme on simulated soil moisture memory, Clim. Dynam., 44, 1731–1750, doi:10.1007/s00382-014-2221-6, 2015.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J., Heimann, M., Hooker-Stroud, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E., Köstler, C., Mathieu, P.-P., Pinty, B., Reick, C., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., Terwisscha van Scheltinga, A., Vossbeck, M., Widmann, H., and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, J. Geophys. Res.-Biogeo., 118, 1414–1426, doi:10.1002/jgrg.20118, 2013.

Kattge, J., Knorr, W., Raddatz, T., and Wirth, C.: Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models, Global Change Biol., 15, 976–991, doi:10.1111/j.1365-2486.2008.01744.x, 2009.

Knauer, J., Werner, C., and Zaehle, A.: Evaluating stomatal models and their atmospheric drought response in a land surface scheme: A multibiome analysis, J. Geophys. Res.-Biogeo., 120, 1894–1911, doi:10.1002/2015JG003114, 2015.

Knorr, W.: Satellite Remote Sensing and Modelling of the Global CO_2 Exchange of Land Vegetation: A Synthesis Study, Examensarbeit, Max-Planck-Institut für Meteorologie, Hamburg, 1894–1911, 1997.

Knorr, W. and Kattge, E.: Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling, Global Change Biol., 11, 1333–1351, doi:10.1111/j.1365-2486.2005.00977.x, 2005.

Kolari, P., Pumpanen, J., Rannik, U., Ilvesniemi, H., Hari, P., and Berninger, F.: Carbon balance of different aged Scots pine forests in Southern Finland, Global Change Biol., 10, 1106–1119, doi:10.1111/j.1529-8817.2003.00797.x, 2004.

Kolari, P., Kulmala, L., Pumpanen, J., Launiainen, S., Ilvesniemi, H., Hari, P., and Nikinmaa, E.: CO_2 exchange and component CO_2 fluxes of a boreal Scots pine forest, Boreal Environ. Res., 14, 761–783, 2009.

Louis, J.-F.: A parametric model of vertical eddy fluxes in the atmosphere, Bound.-Lay. Meteorol., 17, 187–202, 1979.

Mammarella, I., Launiainen, S., Gronholm, T., Keronen, P., Pumpanen, J., Rannik, U., and Vesala, T.: Relative Humidity Effect on the High-Frequency Attenuation of Water Vapor Flux Measured by a Closed-Path Eddy Covariance System, J. Atmos. Ocean. Tech., 26, 1856–1866, doi:10.1175/2009JTECHA1179.1, 2009.

Matheny, A., Bohrer, G., Stoy, P., Baker, I., Black, A., Desai, A., Dietze, M., Gough, C., Ivanov, V., Jassal, R., Novick, K., Schäfer, K., and Verbeeck, H.: Characterizing the diurnal patterns of errors in the prediction of evapotranspiration by several land-surface models: An NACP analysis, J. Geophys. Res.-Biogeo., 119, 1458–1473, doi:10.1002/2014JG002623, 2014.

Morris, M.: Factorial Sampling Plans for Preliminary Computational Experiments, Technometrics, 33, 161–174, 1991.

Mueller, B. and Seneviratne, S.: Systematic land climate and evapotranspiration biases in CMIP5 simulations, Geophys. Res. Lett., 41, 128–134, doi:10.1002/2013GL058055, 2014.

Murphy, J., Sexton, D., Barnett, D., Jones, G., Webb, M., Collins, M., and Stainforth, D.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768–772, doi:10.1038/nature02771, 2004.

Peltoniemi, M., Markkanen, T., Härkönen, S., Muukkonen, P., Thum, T., Aalto, T., and Mäkelä, A.: Consistent estimates of gross primary production of Finnish forests – comparison of estimates of two process models, Boreal Environ. Res., 20, 196–212, 2015a.

Peltoniemi, M., Pulkkinen, M., Aurela, M., Pumpanen, J., Kolari, P., and Mäkelä, A.: A semi-empirical model of boreal-forest gross primary production, evapotranspiration, and soil water – calibration and sensitivity analysis, Boreal Environ. Res.,, 20, 151–171, 2015b.

Rannik, Ü., Peltola, O., and Mammarella, I.: Random uncertainties of flux measurements by the eddy covariance technique, Atmos. Meas. Tech., 9, 5163–5181, doi:10.5194/amt-9-5163-2016, 2016.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, Global Change Biol., 11, 1424–1439, doi:10.1111/j.1365-2486.2005.001002.x, 2005.

Reick, C., Raddatz, T., Brovkin, V., and Gayler, V.: Representation of natural and anthropogenic land cover change in MPI-ESM, J. Adv. Model. Earth Syst., 5, 1–24, doi:10.1002/jame.20022, 2013.

re3data.org: FLUXNET; editing status 2015-05-19; re3data.org – Registry of Research Data Repositories, doi:10.17616/R36K9X, 2016.

Richardson, A., Hollinger, D., Burba, G., Davis, K., Flanagan, L., Katul, G., Munger, J., Ricciutio, D., Stoy, P., Suyker, A., Verma, S., and Wofsy, S.: A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes, Agr. Forest Meteorol., 136, 1–18, doi:10.1016/j.agrformet.2006.01.007, 2006.

Richardson, A., Mahecha, M., Falge, E., Kattge, J., Moffat, A., Papale, D., Reichstein, M., Stauch, V., Braswell, B., Churkina, G., Kruijt, B., and Hollinger, D.: Statistical properties of random CO_2 flux measurement uncertainty inferred from model residuals, Agr. Forest Meteorol., 148, 38–50, doi:10.1016/j.agrformet.2007.09.001, 2008.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM5. PART I: Model description, Max Planck Institute for Meteorology Report 349, 1–127, http://www.mpimet.mpg.de/fileadmin/publikationen/Reports/max_scirep_349.pdf (last access: 7 December 2016), 2003.

Santaren, D., Peylin, P., Bacour, C., Ciais, P., and Longdoz, B.: Ecosystem model optimization using in situ flux observations: benefit of Monte Carlo versus variational schemes and analyses of the year-to-year model performances, Biogeosciences, 11, 7137–7158, doi:10.5194/bg-11-7137-2014, 2014.

Scharnagl, B., Vrugt, J., Vereecken, H., and Herbst, M.: Inverse modelling of in situ soil water dynamics: investigating the effect of different prior distributions of the soil hydraulic parameters, Hydrol. Earth Syst. Sci., 15, 3043–3059, doi:10.5194/hess-15-3043-2011, 2011.

Schulze, E., Kelliher, F., Korner, C., Lloyd, J., and Leuning, R.: Relationships among Maximum Stomatal Conductance, Ecosystem Surface Conductance, Carbon Assimilation Rate, and Plant Nitrogen Nutrition: A Global Ecology Scaling Exercise, Annu. Rev. Ecol. Syst., 25, 629–662, 1994.

Scott, D. W.: Multivariate Density Estimation and Visualization, http://EconPapers.repec.org/RePEc:zbw:caseps:200416 (last access: 7 December 2016), 2004.

Sellers, P.: Canopy reflectance, photosynthesis and transpiration, Int. J. Remote Sens., 6, 1335–1372, doi:10.1080/01431168508948283, 1985.

Singh, R., Reager, J., Miller, N., and Famiglietti, J.: Toward hyper-resolution land-surface modeling: The effects of fine-scale topography and soil texture on CLM4.0 simulations over the Southwestern U.S., Water Resour. Res., 51, 2648–2667, doi:10.1002/2014WR015686, 2015.

Solonen, A., Ollinaho, P., Laine, M., Haario, H., Tamminen, J., and Järvinen, H.: Efficient MCMC for Climate Model Parameter Estimation: Parallel Adaptive Chains and Early Rejection, Bayesian Anal., 7, 715–736, doi:10.1214/12-BA724, 2012.

Suni, T., Rinne, J., Reissell, A., Altimir, N., Keronen, P., Rannik, U., Dal Maso, M., Kulmala, M., and Vesala, T.: Longterm measurements of surface fluxes above a Scots pine forest in Hyytiälä, southern Finland, 1996–2001, Boreal Environ. Res., 8, 287–301, 2003.

Tesfa, T., Li, H.-Y., Leung, L., Huang, M., Ke, Y., Sun, Y., and Liu, Y.: A subbasin-based framework to represent land surface processes in an Earth system model, Geosci. Model Dev., 7, 947–963, doi:10.5194/gmd-7-947-2014, 2014.

Thum, T., Aalto, T., Laurila, T., Aurela, M., Kolari, P., and Hari, P.: Parametrization of two photosynthesis models at the canopy scale in northern boreal Scots pine forest, Tellus B, 59, 874–890, doi:10.1111/j.1600-0889.2007.00305.x, 2007.

Thum, T., Aalto, T., Laurila, T., Aurela, M., Lindroth, A., and Vesala, T.: Assessing seasonality of biochemical CO_2 exchange model parameters from micrometeorological flux observations at boreal coniferous forest, Biogeosciences, 5, 1625–1639, doi:10.5194/bg-5-1625-2008, 2008.

Trusilova, K., Trembath, J., and Churkina, G.: Parameter Estimation and Validation of the Terrestrial Ecosystem Model BIOME-BGC Using Eddy-Covariance Flux Measurements, http://EconPapers.repec.org/RePEc:zbw:caseps:200416 (last access: 7 December 2016), 2004.

Vesala, T., Suni, T., Rannik, Ü., Keronen, P., Markkanen, T., Sevanto, S., Grönholm, T., Smolander, S., Kulmala, M., Ilvesniemi, H., Ojansuu, R., Uotila, A., Levula, J., Mäkelä, A., Pumpanen, J., Kolari, P., Kulmala, L., Altimir, N., Berninger, F., Nikinmaa, E., and Hari, P.: Effect of thinning on surface fluxes in a boreal forest, Global Biogeochem. Cy., 19, GB2001, doi:10.1029/2004GB002316, 2005.

Wu, S., Jansson, P., and Kolari, P.: Modeling seasonal course of carbon fluxes and evapotranspiration in response to low temperature and moisture in a boreal Scots pine ecosystem, Ecol. Model., 222, 3103–3119, doi:10.1016/j.ecolmodel.2011.05.023, 2011.

Paper II

Parameter calibration and stomatal conductance formulation comparison for boreal forests with adaptive population importance sampler in the land surface model JSBACH

Mäkelä, J., Knauer, J., Aurela, M., Black, A., Heimann, M., Kobayashi, H., Lohila, A., Mammarella, I., Margolis, H., Markkanen, T., Susiluoto, J., Thum, T., Viskari, T., Zaehle, S., and Aalto, T.

Geosci. Model Dev., 12, 4075–4095, doi:10.5194/gmd-12-4075-2019, 2019.

II

©Authors 2019.
Reprinted under Creative Commons Attribution 4.0 License.

Geosci. Model Dev., 12, 4075–4098, 2019
https://doi.org/10.5194/gmd-12-4075-2019
© Author(s) 2019. This work is distributed under
the Creative Commons Attribution 4.0 License.

Geoscientific Model Development
Open Access
EGU

Parameter calibration and stomatal conductance formulation comparison for boreal forests with adaptive population importance sampler in the land surface model JSBACH

Jarmo Mäkelä[1], Jürgen Knauer[2], Mika Aurela[1], Andrew Black[3], Martin Heimann[2], Hideki Kobayashi[4], Annalea Lohila[1,5], Ivan Mammarella[5], Hank Margolis[6], Tiina Markkanen[1], Jouni Susiluoto[1,7], Tea Thum[2], Toni Viskari[1], Sönke Zaehle[2], and Tuula Aalto[1]

[1]Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland
[2]Max Planck Institute for Biogeochemistry, 07745 Jena, Germany
[3]University of British Columbia, Vancouver, Canada
[4]Institute of Arctic Climate and Environment Change Research,
Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan
[5]Institute for Atmospheric and Earth System Research/Physics, P.O. Box 48, Faculty of Science,
00014 University of Helsinki, Helsinki, Finland
[6]Department of Forest Sciences, Laval University, Québec city, Canada
[7]School of Engineering Science, Lappeenranta-Lahti University of Technology, P.O. Box 20, 53851 Lappeenranta, Finland

Correspondence: Jarmo Mäkelä (jarmo.makela@fmi.fi)

Received: 7 December 2018 – Discussion started: 6 February 2019
Revised: 16 August 2019 – Accepted: 26 August 2019 – Published: 20 September 2019

Abstract. We calibrated the JSBACH model with six different stomatal conductance formulations using measurements from 10 FLUXNET coniferous evergreen sites in the boreal zone. The parameter posterior distributions were generated by the adaptive population importance sampler (APIS); then the optimal values were estimated by a simple stochastic optimisation algorithm. The model was constrained with in situ observations of evapotranspiration (ET) and gross primary production (GPP). We identified the key parameters in the calibration process. These parameters control the soil moisture stress function and the overall rate of carbon fixation.

The JSBACH model was also modified to use a delayed effect of temperature for photosynthetic activity in spring. This modification enabled the model to correctly reproduce the springtime increase in GPP for all conifer sites used in this study. Overall, the calibration and model modifications improved the coefficient of determination and the model bias for GPP with all stomatal conductance formulations. However, only the coefficient of determination was clearly improved for ET. The optimisation resulted in best performance by the Bethy, Ball–Berry, and the Friend and Kiang stomatal conductance models.

We also optimised the model during a drought event at a Finnish Scots pine forest site. This optimisation improved the model behaviour but resulted in significant changes to the parameter values except for the unified stomatal optimisation model (USO). Interestingly, the USO demonstrated the best performance during this event.

1 Introduction

Plants exchange carbon dioxide (CO_2) and water vapour (H_2O) with the atmosphere. Sufficient soil water, irradiance and adequate temperature are required to maintain the exchange rates during the growing season. Disturbances in these conditions such as drought, cold temperature or low radiation cause the plants to respond to the environmental stress via stomatal closure and the decrease in photosynthesis and transpiration (Lagergren and Lindroth, 2002; Mäkelä et al., 2004; Gao et al., 2017). The capability of plants to recover from such events depends on species and their adaptation to site conditions (Kozlowski and Pallardy, 2002). Stress

is part of the normal annual cycle of the plants, but occasionally it may exceed the limits of recovery.

Soil water deficit and high water vapour pressure deficit can result in suppressed plant transpiration (Bréda et al., 1993; Kropp et al., 2017). Globally, soil drought has been recognised as one of the main limiting factors for plant photosynthesis (Nemani et al., 2003), and boreal forests are known to occasionally suffer from soil drought (Muukkonen et al., 2015; Gao et al., 2016). The recovery of photosynthetic capacity in spring has been connected to temperature history and to the frequency of severe night frosts (Bergh et al., 1998; Bergh and Linder, 1999), which can reverse the recovery. Understanding and correctly modelling these phenomena are especially important for boreal forests (Bonan, 2008) under changing environmental conditions.

Ecosystem and land surface models, describing the plant photosynthesis, transpiration and soil-hydrology-related processes, usually include descriptions and parameterisations for various stress effects. These parameters often lack a theoretical foundation (Gao et al., 2002; Medlyn et al., 2011), and descriptions of vegetation drought response and phenology have been recognised to need better formulations and design (Richardson et al., 2012; Powell et al., 2013; Xu et al., 2013; Medlyn et al., 2016). These deficiencies restrict a model's predictive capability under changing environmental conditions and call for specific parameterisations for different plant types and vegetation zones.

Stomatal conductance models describe the pathway of CO_2 and water through the leaf stomata by an electric circuit analogy (Nobel, 1999). The variations in stomatal opening and mesophyll structure are interpreted as resistances to water flow and the process is idealised via generalised parameterisation. Stomatal conductance models mainly differ in their choice of variables driving the stomatal closure, and their performance has recently been assessed in modelling studies by, e.g., Egea et al. (2011), Knauer et al. (2015) and Franks et al. (2018). However, it can be hypothesised that the choice of the stomatal conductance model affects the ecosystem model parameters more broadly as the stomatal conductance formulations vary in their responses to the different conditions. A holistic assessment of the performance of the stomatal conductance models together with ecosystem model parameter optimisation has been missing.

In many other studies, where the aim has been to optimise land surface model parameters, the optimisation is based on estimating the gradient of the cost function: Knorr et al. (2010) for JSBACH, Kuppel et al. (2012) and Peylin et al. (2016) for ORCHIDEE, and Raoult et al. (2016) for JULES. Gradient-based methods are faster than Markov chain Monte Carlo (MCMC) methods as they strongly steer the sampling process to reach a minimum in the cost function (see, e.g., Gelman et al., 2013). This approach also enables a more indefinite setting of parameter ranges (limits for acceptable parameter values) when compared to methods that sample the full parameter space. However, they are prone to get stuck in local minima, especially when the dimensionality of the parameter space increases. In the last few years, similar parameter estimations have also been done for CLM (Community Land Model) by Post et al. (2017) using the $\mathrm{DREAM}_{(\mathrm{zs})}$ (MCMC) algorithm with multiple chains and for JULES by Iwema et al. (2017) with the BORG algorithm that employs multiple optimisation algorithms simultaneously. The DREAM algorithm is fully iterative, which limits the number of parallel processes to the number of parallel chains in use (when we do not account for the possibility of the model parallelisation, that can be substantial). The applicability of the BORG algorithm is dependent on the algorithms in use and the expertise of the user (to choose the right algorithms, etc.).

APIS (adaptive population importance sampler) is a Monte Carlo (MC) method that can be run iteratively as presented by Martino et al. (2015), but it is also straightforward to parallelise, since all samples prior to each adaptation (in our simulation 2000 draws) can be drawn and estimated simultaneously. This latter feature is useful to decrease the amount of real time required to run the algorithm when computer resources are not the limiting factor – APIS requires considerably fewer sequential estimates than typical Markov chain methods. In the iterative mode, automatic stopping rules can be easily implemented to indicate when additional samples are not required to improve the estimates. The APIS algorithm samples the full parameter space (as do MCMC methods) and can utilise a mixture of parameter prior distributions. Therefore, APIS can estimate complicated multidimensional probability distributions with relative ease. These aspects make APIS an attractive alternative to the other sampling and optimisation methods mentioned above.

In this study we apply the land surface model JSBACH for 10 boreal coniferous evergreen forest eddy covariance sites to examine the performance of different stomatal conductance models, and their effect on calibrated parameters related to photosynthesis, phenology and hydrology. First, we utilise APIS to sample the full parameter space with the different stomatal conductance formulations and to locate different modes of the target distributions (peaks of high probability). Second, using the distributions generated by APIS as the prior distributions, we optimise the parameters using a simple stochastic optimisation method. Finally, we assess the inter-site variability and the robustness of the calibrated parameters together with different stomatal conductance formulations. Optimised parameters for a specific drought are also investigated and compared with the parameters for the general optimisation.

2 Materials and methods

We will next introduce the measurement sites, followed by the model and modifications made to it. Afterwards we will give a general overview of the simulations as well as the sam-

pling process, the algorithms and methods used to analyse the results.

2.1 Sites and measurements

We use data from 10 FLUXNET (Baldocchi et al., 2011) sites characterised as coniferous evergreen forests. Site descriptions with appropriate references are provided in Table 1. The site-level half-hourly eddy covariance (EC) measurements were quality checked and gap-filled when needed to produce continuous half-hourly and daily time series. The gap-filled and low-quality (based on FLUXNET data quality flags) measurements were masked, and the daily aggregates (usually means) were accepted as part of the calibration process if at least 60 % of the values between 04:00 and 20:00 (i.e. daytime measurements) for that day were unmasked. The daily aggregates of evapotranspiration (ET) and gross primary production (GPP) were used to calibrate and validate the model, whereas the half-hourly data were used as climate forcing (as explained later in Sect. 2.4.

Based on the quality and quantity of their respective measurements, the sites were divided into calibration and validation sites. Essentially, if we have enough data from a site, it is used for both calibration and validation purposes. We required the site to have at least 8 years of measurements, where the first five were used for calibration, and the consecutive three for validation. Otherwise we used the site only for a 3-year validation. The FLUXNET datasets were missing both the long- and shortwave radiation for the two Russian sites – Fyodorovskoye (RU-Fyo) and Zotino (RU-Zot). These were generated from ERA Interim data. The soil types of all of these sites can mostly be identified as mineral soils with varying sand, clay and peat contents. Fyodorovskoye and Poker Flat (US-Prr) are natural peatlands, and Lettosuo (FI-Let) is a drained peatland site.

The measurement error in the EC flux data was separated into systematic and random errors. The main systematic errors (density fluctuations, high-frequency losses, calibration issues) were taken into account as part of the post-processing of the data, and the random errors tend to dominate the uncertainty of the instantaneous fluxes. The random error is often assumed to be Gaussian but can be more accurately approximated by a symmetric exponential distribution (Richardson et al., 2006). It increases linearly with the magnitude of the flux, with a standard deviation typically less than 20 % of the flux (Richardson et al., 2008; Rannik et al., 2016). Our treatment of the measurement (and model) errors is explained in Sect. 2.9.

2.2 The JSBACH model

JSBACH (Kaminski et al., 2013) is a process-based ecosystem model and the land surface component of the Earth System model of the Max Planck Institute for Meteorology (MPI-ESM). We ran JSBACH offline using meteorological measurements from the flux towers to force the model. Implications of this one-way coupling with the atmosphere include lack of feedback from the surface energy balance to the atmosphere; i.e. latent and sensible heat fluxes and surface thermal radiation do not directly affect prescribed air temperature or humidity. Similarly, the feedback of the surface to the vertical transfer coefficients within the atmospheric surface layer is missing as the wind speed that drives mixing is prescribed. Furthermore, since we use site-level data (each site is represented as a single grid point), the grid resolution does not affect the results.

We focus only on the most essential parts of JSBACH relating to our work. A more complete model description with details on, e.g., soil heat transfer, water balance and coupling to the atmosphere can be found in Roeckner et al. (2003), whereas Raddatz et al. (2007) provides a more descriptive synopsis on land–surface interactions, Reick et al. (2013) complements both with an addition of land cover change processes, and Hagemann and Stacke (2015) introduces soil hydrological mechanisms within a multilayer scheme applying five layers.

In JSBACH, the land surface is divided into grid cells, which are split into bare soil and vegetative areas. The vegetative area is further divided into tiles representing the most prevalent vegetation classes, called plant functional types (PFTs) (Reick et al., 2013). In our site-level simulations, the model was set to use only one PFT: coniferous evergreen trees. The seasonal development of leaf area index (LAI) for the trees is regulated by air temperature and soil moisture with a single limiting value (for all sites) for the maximum of LAI. This maximum value was fixed and the site-specific fractions of vegetative area were adjusted to reproduce the measured site-level LAI.

The predictions of phenology are produced by the Logistic Growth Phenology (LoGro-P) sub-model in JSBACH (Böttcher et al., 2016). Photosynthesis is described by the biochemical photosynthesis model (Farquhar et al., 1980). Following Kattge et al. (2009), we set the maximum electron transport rate (J_{max}) at 25 °C to 1.9 times the maximum carboxylation rate ($V_{C,max}$), which is in line with, e.g., Leuning (2002) and Ueyama et al. (2016). The photosynthetic rate is dependent on the stomatal conductance formulation used, introduced in Sect. 2.3. Radiation absorption is estimated by a two stream approximation within a three-layer canopy (Sellers, 1985). Especially in sparse canopies, radiation absorption is affected by clumping of the leaves which is here taken into account according to the formulation by Knorr (1997).

Parameters detailing site-specific soil properties, such as soil porosity and field capacity, were derived from FLUXNET datasets and the references in Table 1. We approximated the soil composition and generated these properties following Hagemann and Stacke (2015).

Table 1. Descriptions for the sites used in this study sorted by their FLUXNET identifier. The first six sites are used for both calibration and validation purposes, with the first 5 years of each site used for calibration. The last 3 years as well as the last four sites are used for validation only. The reported elevation is in metres above sea level, LAI is the one-sided leaf area index, and the average stand age is in years, along with average annual precipitation (P) in millimetres and temperature (T) in degrees Celsius.

Site id	Lat	Long	Elev.	Dom. species	LAI	Age	P	T	Years	Reference
CA-Obs	53.99	−105.12	629	*Picea mariana*	3.8	135	406	0.8	1999–2006	Chen et al. (2006)
CA-Qfo	49.69	−74.34	382	*Picea mariana*	3.7	112	962	−0.4	2003–2010	Chen et al. (2006)
FI-Hyy	61.85	24.29	180	*Pinus sylvestris*	3.5	45	709	2.9	1999–2006	Kolari et al. (2009)
FI-Ken	67.99	24.24	337	*Picea abies*	2.1	100	484	0.4	2003–2010	Aurela et al. (2015)
FI-Sod	67.36	26.64	179	*Pinus sylvestris*	1.7	150	527	−0.4	2001–2008	Thum et al. (2007)
RU-Fyo	56.45	32.90	265	*Picea abies*	4.5	200	711	3.9	2002–2009	Launiainen et al. (2016)
CA-Ojp	53.92	−104.69	579	*Pinus banksiana*	2.6	100	431	0.1	2004–2006	Chen et al. (2006)
FI-Let	60.64	23.96	119	*Pinus sylvestris*	6.0	40	627	4.6	2010–2012	Launiainen et al. (2016)
RU-Zot	60.80	89.35	121	*Pinus sylvestris*	1.5	215	493	−3.3	2002–2004	Kelliher et al. (1998)
US-Prr	65.12	−147.49	210	*Picea mariana*	0.7	72	275	−2.0	2011–2013	Ikawa et al. (2015)

2.3 Modifications to the JSBACH model

All parameters of interest, presented in Table 2, were extracted from the JSBACH model code to an external file to facilitate the simulations. The default values of newly added parameters (not originally in JSBACH: τ, q, g_0, g_1) were derived from a synthesis of literature values. Most of the parameter ranges (limiting values for the parameters) were adapted from our previous work on a similar topic (Mäkelä et al., 2016). The parameter grouping was done to enhance optimisation, and the mechanism is explained in Sect. 2.7. Group I consists of parameters most directly affecting photosynthesis, group II parameters are intimately involved with soil moisture, and group III are the LoGro-P model parameters. The equations governed by these parameters are presented in Appendix A.

The start of the growing season in the JSBACH model is defined by a "spring event" in LoGro-P (Appendix A3) that induces leaf growth. The phenology model calculates a sum of ambient temperature (heatsum) since last autumn that is above the cutoff value T_{alt}, presented in Eq. (A10). It also calculates a variable threshold, defined in Eq. (A12), for the heatsum to reach. The threshold decreases based on the number of days the ambient temperature is below T_{alt}, whereas the heatsum increases. When the heatsum reaches the threshold, the plant leaves are free to grow.

However, coniferous evergreen trees do not shed all of their leaves for winter and the existing foliage enables them to quickly initiate photosynthesis in the following spring. The start of the photosynthetically active season in the model has been observed to occur too early in the boreal region by, e.g., Böttcher et al. (2016). In order to correct this behaviour, i.e. to restrain the respiration and photosynthesis of conifers in the early spring, we utilise a delayed effect of temperature for photosynthetic activity, introduced by Mäkelä et al. (2004). To calculate the reduction, we must first define the state of photosynthetic acclimation that Mäkelä et al. (2004, p. 371) present as "an aggregated measure of the state of those physiological processes of the leaves that determine the current photosynthetic capacity at any moment".

The state of acclimation (S) is calculated from air temperature (T) with a delay prescribed by parameter τ (this is similar to the calculation of T_{S} in Appendix A14). S is then inserted into sigmoidal relation Eq. (1) to calculate a factor γ, a formulation that is adapted here from Kolari et al. (2007). Finally, γ is used to reduce the photosynthetic efficiency in Eq. (A1). $T_{1/2}$ denotes the inflection point where γ reaches half of $\gamma_{\max}$, k is the curvature of the function and $\gamma = 1$ when $S \geq 10$.

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{T - S}{\tau}, \gamma = \frac{\gamma_{\max}}{1 + e^{k(S - T_{1/2})}} \tag{1}$$

The JSBACH model was also modified to include altogether six different stomatal conductance formulations following Knauer et al. (2015). These formulations include the pre-existing Baseline and Bethy versions as well as the Ball–Berry model and three of its variants. Model information is gathered in Table 3 for easy referencing and the detailed formulations are given in Appendix B. The limits of the slope of the stomatal conductance formulation parameter (g_1) were set to reflect commonly observed values from physiological measurements (Egea et al., 2011). The limits of g_1^{USO} reflect the results presented by Lin et al. (2015).

We have also included two additional parameters (a and d in Table 2) for the Friend and Kiang (Friend and Kiang, 2005) stomatal conductance formulation in Eq. (B3). These parameters were not originally included in the optimisation, but the resulting cost function (Eq. 9) values were poor when compared to the other formulations. At that point, these parameters were included in the optimisation process. This increases the degrees of freedom for the Friend and Kiang model by two and therefore may give it an advantage when compared to the other Ball–Berry type formulations, which has to be considered in the interpretation of the results.

Table 2. Descriptions of model parameters with default values, range of acceptable values and references to equations in the paper or in the appendices. Parameters in the same group were calibrated simultaneously.

Parameter	Def.	Range	Units	Group	Description	Eq.
$V_{C,max}$	62.5	[40, 65]	◇	I	Farquhar model maximum carboxylation rate at 25 °C of the enzyme Rubisco (coupled with maximum electron transport rate at 25 °C with a factor of 1.9) (◇ = µmol (CO_2) m^{-2} s^{-1}).	(A2)
α	0.28	[0.26, 0.32]	–	I	Farquhar model efficiency for photon capture at 25 °C.	(A4)
τ	10.0	[5, 15]	d	I	Adjustment period length in acclimation of photosynthesis.	(1)
c_b	5.0	[4, 7]	–	I	Multiplier in momentum and heat stability functions (Louis, 1979).	–
f_{C3}	0.87	[0.7, 0.95]	–	I	Ratio of unstressed C3-plant internal/external CO_2 concentration.	(A3)
q	0.0	[0,1]	–	I	Exponential scaling of water stress in reducing photosynthesis.	(A1)
g_0	0.001	$[1 \times 10^{-5}, 5 \times 10^{-3}]$	▽	I	Residual stomatal conductance (▽ = mol m^{-2} s^{-1}).	(B3)
g_1	Values in Table (3)		–	I	Slope of the stomatal conductance function.	(B3)
a	2.8	[1.5, 3.5]	–	I	Base rate of stomatal conductance response to atmospheric humidity for the Friend and Kiang model.	(B3)
d	80	[50, 120]	–	I	Exponential rate of stomatal conductance response to atmospheric humidity for the Friend and Kiang model.	(B3)
θ_{dr}	0.9	[0.5, 0.95]	–	II	Volumetric soil water content above which fast drainage occurs.	(A6)
θ_{hum}	0.5	[0.2, 0.8]	–	II	Fraction depicting relative surface humidity based on soil dryness.	(A9)
θ_{pwp}	0.35	[0.15, 0.4]	–	II	Volumetric soil moisture content at permanent wilting point.	(2)
θ_{tsp}	0.75	[0.25, 0.8]	–	II	Value of volumetric soil moisture content above which transpiration is unaffected by soil moisture stress (β); $0.9\theta_{tsp} \geq \theta_{pwp}$.	(2)
p_{int}	0.25	[0.15, 0.35]	–	II	Fraction of precipitation intercepted by the canopy.	(A5)
s_{sm}	5.9×10^{-3}	$[1 \times 10^{-4}, 0.1]$	m	II	Depth for correction of surface temperature for snow melt.	–
w_{skin}	2.0×10^{-4}	$[1 \times 10^{-5}, 5 \times 10^{-3}]$	m	II	Maximum water content of the skin reservoir of bare soil.	–
C_{decay}	13.0	[5, 25]	d	III	LoGro-P: memory loss parameter for chill days.	(A12)
S_{min}	10.0	[5, 30]	°C d	III	LoGro-P: minimum value of critical heat sum.	(A12)
S_{range}	150.0	[100, 300]	°C d	III	LoGro-P: maximal range of critical heat sum.	(A12)
T_{alt}	4.0	[2, 10]	°C	III	LoGro-P: cutoff in alternating temperature.	(A10)
T_{ps}	10.0	[3, 25]	°C	III	LoGro-P: memory loss parameter for pseudo soil temperature.	(A14)

All of the stomatal conductance models contain an empirical water stress factor β, which reduces stomatal conductance as a function of volumetric soil water content (θ).

$$\beta = \begin{cases} 1, & \theta \geq \theta_{tsp} \\ \frac{\theta - \theta_{pwp}}{\theta_{tsp} - \theta_{pwp}}, & \theta_{pwp} < \theta < \theta_{tsp} \\ 0, & \theta \leq \theta_{pwp} \end{cases} \quad (2)$$

In JSBACH, the stomatal conductance (g_s) is primarily resolved to estimate carbon fixation. The same g_s is then later used to calculate transpiration (Eq. A8). In the original JSBACH formulation (i.e. the Baseline version), the g_s is first resolved for unstressed canopy and then scaled by the water stress factor β. The Bethy approach is similar, but the conductance can also be limited by water supply (Eq. B2). In cases when the water supply is not the limiting factor, the calculations are similar to the Baseline version. In all of the empirical Ball–Berry variants, the stomatal conductance can be written as $g_s = g_0 + c\beta g_1$. The residual conductance (g_0) and the slope of the function (g_1) are both formulation-specific parameters as well as the factor c, which incorporates net photosynthesis and effects of atmospheric humidity and CO_2 concentration. The parameters g_0 and g_1 are part of our sampling and optimisation processes (group I in Table 2 when applicable).

The water stress factor (β) limits the carbon fixation and transpiration via the stomatal conductance formulation. Following Egea et al. (2011), it is also used to directly limit the net assimilation rate (A_n), as seen in Eq. (A1). The additional scaling (or limiting) factor for A_n takes the form β^q, so it is a function of both soil water content θ and the parameter q.

Maximal reduction is achieved when $q = 1$, and the reduction factor reverts to β. The minimal reduction occurs when $q = 0$ and the reduction factor resembles a step function (at $\theta = \theta_{\mathrm{pwp}}$). For any other value of q, it is a continuous convex function between the two extremes $\beta^q : [\theta_{\mathrm{pwp}}, \theta_{\mathrm{tsp}}] \rightarrow [0, 1]$.

2.4 Model simulations

The site-level measurements, used as model inputs, are air temperature, air pressure, precipitation, humidity, wind speed and CO_2 concentration as well as short- and longwave and potential shortwave radiation. Additionally, ET and GPP, derived from the EC measurements, are used to constrain and evaluate the model (as explained later in Sect. 2.8 and 2.9). We drive the model with half-hourly data but output daily values.

The initial state of the JSBACH model can be generated from predefined values of state variables (usually empty initial storage pools) or the model can be restarted from a file describing the state of some previous run. Depending on the area of interest, a model spin-up may be required to bring the model into a steady state. In our simulations, some of the more slowly changing variables (e.g. soil water content and LAI) need to be equilibrated, so a spin-up is required. This can be achieved by running the model over a set of measurements multiple times, each time restarting from the final state of the previous run.

The calibration period consists of the first 5 years given for the calibration sites in Table 1. The spin-up is achieved by looping over these 5 years, altogether four times (20-year spin-up), and then saving the state of the model at the end of the run. The actual calibration is started from the beginning of the calibration period, using the previously saved state variables. To reduce any bias this induces, the first year in the calibration run is removed from the cost function calculations. The spin-ups for the validation sites in Table 1 are similarly generated.

During the summer 2006, the Hyytiälä (FI-Hyy) measurement site suffered from a severe drought (Gao et al., 2017), leading to visible discolouration of needles. These events are difficult for models to capture and hence are of interest to modellers. We have previously and unsuccessfully attempted to optimise the JSBACH model (Mäkelä et al., 2016) for this event. Here we focus directly on the extended dry period (190–260th day of the year in 2006), during which the actual drought is mostly in effect between 210 and 235th DOY (day of the year). We adjusted some of the parameter values as those uncovered by the more general calibration, presented above. The spin-up was the same as for the calibration period, but at the end of the spin-up, the model was run forward to the start of the year 2006. Only values between the 190 and 260th DOY in 2006 were used in constraining the model.

2.5 Sampling process

We describe the modelling setup with the equation $\mathbf{y} = \mathcal{M}(\boldsymbol{\theta}, \mathbf{x}) + \mathbf{e}$, where the aim is to reproduce the observations ($\mathbf{y}$) with our model ($\mathcal{M}$), the driving data ($\mathbf{x}$) and the current parameter values ($\boldsymbol{\theta}$). The residuals ($\mathbf{e}$) depict how well the model reproduces the observations, and they form the basis of the likelihood function (formulated in Sect. 2.9), which is used to derive the parameter posterior distributions.

Using Bayes' rule on conditional probability we can write the parameter posterior density ($p(\boldsymbol{\theta}, \mathcal{M}|\mathbf{x})$) as a function of the likelihood ($\mathcal{L}(\mathbf{x}|\boldsymbol{\theta}, \mathcal{M})$), parameter prior distributions ($\pi(\boldsymbol{\theta})$) and the model evidence ($Z(\mathbf{x}|\mathcal{M})$). As usual and from here on, we do not write $\mathcal{M}$ in the Bayes' formula:

$$p(\boldsymbol{\theta}|\mathbf{x}) = \frac{\mathcal{L}(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{Z(\mathbf{x})}. \tag{3}$$

We can now utilise the posterior density as a probability density function (pdf) for the parameters and infer the expectation values:

$$E[\boldsymbol{\theta}_i] = \frac{1}{Z}\int \boldsymbol{\theta}_i\, p(\boldsymbol{\theta}|\mathbf{x})\mathrm{d}\boldsymbol{\theta}, \quad Z = \int p(\boldsymbol{\theta}|\mathbf{x})\mathrm{d}\boldsymbol{\theta}. \tag{4}$$

Above $\boldsymbol{\theta}_i$ is the ith element of the parameter vector. Generally, Eq. (4) cannot be analytically solved; hence it is usually estimated numerically. Commonly this is achieved by one of the many MCMC methods, but in this study we apply APIS defined by Martino et al. (2015). APIS (Martino et al., 2015) is a Monte Carlo (MC) method that utilises a population of importance samplers (IS) to jointly estimate the target pdf ($p(\boldsymbol{\theta}|\mathbf{x})$) and the normalising constant ($Z(\mathbf{x})$) by a deterministic mixture approach (Veach and Guibas, 1995; Owen and Yi, 2000), whereas the MCMC methods do not care about the value of Z. We denote the importance sampling density as $q(\boldsymbol{\theta})$.

$$E[\boldsymbol{\theta}_i] = \frac{1}{Z}\int \boldsymbol{\theta}_i r(\boldsymbol{\theta}) q(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}, \text{ where } r(\boldsymbol{\theta}) = \frac{p(\boldsymbol{\theta}|\mathbf{x})}{q(\boldsymbol{\theta})} \tag{5}$$

Above r is the reweighing factor that is the driving force in importance sampling. We will next give a summary description of the sampling process with comparison to a general multichain MCMC approach (since MCMC methods are more commonly used in these types of situations).

1. The initialisation of a multichain MCMC sampler and APIS are very similar. In our simulations, APIS is set up as 40 simultaneous and independent importance samplers. This is similar to an independent 40-chain MCMC sampler. Each sampler or chain has a random starting location drawn from a uniform distribution defined by the parameter ranges, given in Table 2. The initial sampling (or prior) distribution for each sampler is also randomly generated – we use truncated Gaussian distributions with diagonal covariance matrices, where the standard deviations are randomised. The sampling distributions will evolve throughout the process.

Table 3. Stomatal conductance models with default values and range for g_1 and references to equations in Appendix B as well as related articles.

Stomatal conductance model	Short	g_1	Range	References	
Baseline	Base	–	–	B1	Knorr (1997)
Biosphere–Energy–Transport–Hydrology	Bethy	–	–	B2	Knorr (2000)
* Ball–Berry	BB	9.0	[4,10]	B3	Ball et al. (1987)
* Leuning	Leu	8.0	[6,10]	B3	Leuning (1995)
* Friend and Kiang	F&K	9.5	[7,11]	B3	Friend and Kiang (2005)
* Unified stomatal optimisation	USO	2.0	[1.5,3.5]	B3	Medlyn et al. (2011)

The * symbol indicates the Ball–Berry model and its variants.

2. In an MCMC setup, the model would be run once (for each chain) and evaluated and then the draw (parameter values) accepted or rejected accordingly. In APIS, instead of a single element (one run) we use a sample size of 50. This means that we draw 50 elements with each IS sampler (or "chain") independently. These draws are then evaluated and reweighted as presented in Eq. (5).

3. The 50 reweighted draws (for each IS sampler separately) are used to calculate a new location for the sampling distribution. This location is automatically accepted (no rejection criteria), and we also adapt the shape of the distribution using the self-normalising estimator by Cornuet et al. (2012).

4. Additionally, all of the draws in APIS are used to calculate "global" estimates of the parameter expected values. This process utilises the deterministic mixture approach (Veach and Guibas, 1995; Owen and Yi, 2000) and is fully iterative with no need for any recalculations as the previous estimates are directly adjusted (no information is lost either).

MCMC chains track the evolution of single elements and occasionally adjust the sampling distribution. The sample size in APIS is larger (it is not a Markov chain method) and the focus is on the evolution of the locations of the sampling distributions, not on the individually drawn elements. These location parameters are expected to be around all the modes of the target and the deterministic mixture ensures the stability of the estimation of the (global) parameter expected values. As an importance sampler, APIS is also a variance reducing method.

Before taking a more detailed look at APIS, we make some further notes about the sampling process. The first element of the 50 draws (item 2 in the list above) is always fixed as the current mean. We run the spin-up (Sect. 2.4) and generate the model starting state only for the proposal means and use the same state for the other 49 draws (perturbed around the proposal mean). This requirement stems from a need to reduce computational time as running the model to a steady state is costly. This approach might induce some discrepancies, but they are mitigated by removing the first year of the calibration simulations (as explained in Sect. 2.4). We also slightly reduce the importance weights of the 49 samples (more reduction for samples further from the proposal mean), when calculating the new location parameters (item 3 in the list above) – the reduction only (slightly) slows the adaptation of the IS sampler locations. Finally, we note that this approach ensures that we run the proposal means, which are the focus in APIS, with the correct spin-up.

2.6 Adaptive population importance sampler

Normally, only the location parameters of the IS proposals are adapted, but we also adapt the shape parameters using the self-normalising estimators by Cornuet et al. (2012). APIS is able to utilise different or a mixture of normalised proposals densities, but we use truncated Gaussian proposals with diagonal covariance matrices.

In our simulations, APIS is formed of 40 independent IS estimators. Each estimator draws a sample $\boldsymbol{\theta}_i, i \in \{1, \ldots, N\}$, of size $N = 50$ at a time from their own proposal distribution $q_j(\boldsymbol{\theta}), j \in \{1, \ldots, M\}, M = 40$. The estimator then calculates the importance weights ($w_{ij} = \frac{p(\boldsymbol{\theta}_i|\mathbf{x})}{q_j(\boldsymbol{\theta}_i)}$) for each sample. The location ($\boldsymbol{\mu}_j$) and shape ($\mathbf{C}_j$) parameters (Cornuet et al., 2012) of each proposal are updated using only samples (and weights) drawn from q_j. The new shape parameters are formed as a mean of the previous estimate and $\mathbf{C}_j$, as calculated below.

$$\boldsymbol{\mu}_j = \frac{\sum_i w_{ij}\boldsymbol{\theta}_i}{\sum_i w_{ij}}, \quad \mathbf{C}_j = \frac{\sum_i w_{ij}(\boldsymbol{\theta}_i - \boldsymbol{\mu}_j)(\boldsymbol{\theta}_i - \boldsymbol{\mu}_j)^T}{\sum_i w_{ij}} \tag{6}$$

The simple IS estimators alone are rarely sufficient if the target is even slightly complicated. One classical way of tackling this problem is to join multiple IS estimators together. The simplest approach is to calculate the weights for each of these estimators separately and to normalise the result by the combined sum of all weights. However, this leaves the estimators susceptible to "bad" proposals. APIS suppresses the bad proposals by utilising the deterministic mixture approach (Veach and Guibas, 1995; Owen and Yi, 2000) presented in Eq. (7), where each proposal q_j is evaluated at all

the drawn samples and weighted by the amount of samples drawn ($N_j = 50$) from that proposal. This is equivalent to joining the normalised proposal densities together and evaluating the joint pdf.

$$w_{ij} = \frac{p(\boldsymbol{\theta}_{ij}|\mathbf{x})}{\sum_j \left(\frac{N_j}{\sum_k N_k}\right) q_j(\boldsymbol{\theta}_{ij})} \quad (7)$$

The parameter expectation values and the normalising constant in Eq. (5) can now be estimated by Monte Carlo integration using weights calculated in Eq. (7).

2.7 Parameter optimisation

The APIS algorithm is a rather robust method meant for examining the full target probability distribution and locating the modes of the target distribution. Adaptation in APIS utilises multiple draws simultaneously, which can easily lead to few parameters controlling this process (the marginal density of one or few parameters dominates the calculations). Since we also did not run the model spin-up for all drawn samples (although the discrepancies should be minimal), we utilise a simple custom stochastic optimiser to locate the optimal set of parameter values. This optimiser is run after the APIS calibration simulations and separately for the drought period. The optimiser utilises the exact same datasets (calibration, validation, observations, etc.) as APIS, the spin-up is generated for all drawn samples separately and the initial state of the algorithm is the mean value of the APIS final configuration (location parameters).

Our optimiser is a simple random sampler amplified by the "velocity" of the last jump (the idea is similar to Hamiltonian or Hybrid Monte Carlo by Duane et al., 1987). We draw a set of samples from a small Gaussian proposal distribution in the vicinity of the current best estimate and calculate the cost function for the samples. Whenever a better point is found (smaller cost function), we jump to that (update the mean of the proposal distribution). The velocity of the jump (for us merely the distance of change in each parameter) is then added to the new mean (with a maximal limit of 1 standard deviation in the proposal distribution), but it is reduced and eventually removed if a better sample is not found.

The covariance matrix of the proposal distribution is recalculated at predefined intervals (for all parameters). Additionally, we utilise a subset sampling procedure, where the samples are first drawn from the full parameter space; in the next step they are drawn only from group I in Table 2 (the rest are kept at their current optimal values), followed by groups II and III and then back to the full parameter space. When the number of parameters is reduced, we are more likely to find a better set of parameter values. We have kept the parameters mostly affecting the same processes in the same group, but some dependencies may not be apparent, and hence it is also important to draw samples from the full parameter space.

2.8 Simulation analysis

Even though APIS is not a Markov chain method, we can (naively) interpret the evolution of the location parameters of each IS sampler as chains. The resulting 40 chains have random starting positions, but they are relatively short (we present results from the Bethy calibration, where the chains were adjusted 100 times); hence we did not discard any of the samples. We test the convergence of these chains with the Gelman–Rubin diagnostic tests (Gelman and Rubin, 1992), comparing the variance between the chains to the variance within each chain and calculating the potential scale reduction factors ($\hat{R}$). We also test the stability of the (parameter) global expected value estimate (using the deterministic mixture approach) by calculating the difference of the final global expected value and the mean of the location parameters (at each iteration). We denote this test as δ and report the number of the iterations when this difference is below 5 % of the parameters range, given in Table 2.

In order to visualise the results, we have utilised a Gaussian kernel density estimation (KDE) to produce distributions from the APIS simulation location parameters. In practice, KDE places a Gaussian distribution centred at each sample, and the constructed composite distribution is an estimate of the underlying actual distribution. The bandwidth for the distributions is calculated using Scott's rule (Scott, 2004): the data covariance matrix is multiplied by a factor $n^{\frac{-1}{d+4}}$, where n is the number of data points and d is the number of dimensions.

The effectiveness of each parameter was calculated from the final state of each optimisation process. This was done by first setting all parameters to their optimised values. Then we (evenly) sampled each parameter separately from their range of acceptable values, given in Table 2, and calculated the corresponding cost functions. For each parameter the maximum difference in these cost function values (and the optimised value) was recorded. The parameters (within each optimisation) were then ordered by these numbers (with the highest difference meaning highest effectiveness) and separated into three groups: highest (most effective) and lowest (least effective) effectiveness values and the rest. This effectiveness relates to how the APIS "sees" the sampling process – the 50 draws are evaluated simultaneously, and a very effective parameter can easily mask the influence of a less effective one (the marginal density of one or few parameters dominates the calculations).

We report the slope of the regression line (b) and the coefficient of determination (r^2) between the observations (y_i) and the model output (x_i). The slope of the regression line is highly indicative of the model bias (the difference of the expected values of the observations and the model). Hence we interpret the bias directly from b (in our results the regression lines pass near origin so the differences this induces are negligible).

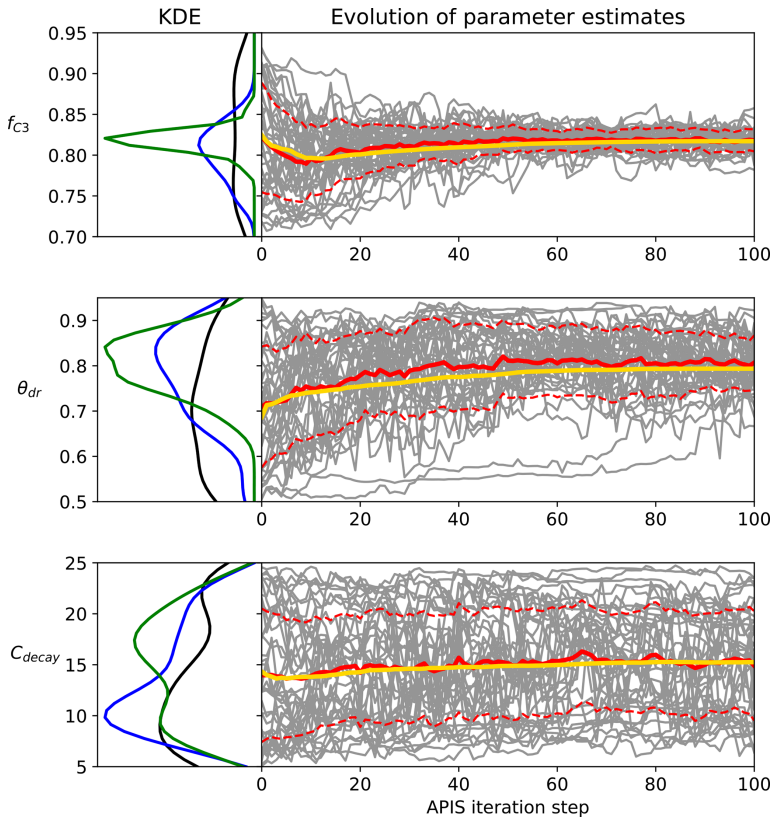

Figure 1. Examples of the evolution of the APIS algorithm from the Bethy calibration. The left panel is the kernel density estimate of the location parameters at the start of the process (black), after 20 iterations (blue) and after 100 iterations (green). The right panel shows the location parameters (grey), their mean (red) and 1 standard deviation (dashed) as well as the global estimate (yellow, calculated with the deterministic mixture approach) of the parameter expected value.

$$b = \frac{\sum_i (x_i - \overline{x_i})(y_i - \overline{y_i})}{\sum_i (y_i - \overline{y_i})^2}, \quad r^2 = 1 - \frac{\sum_i (x_i - y_i)^2}{\sum_i (y_i - \overline{y_i})^2} \tag{8}$$

2.9 Cost function

The Bayesian framework requires a likelihood function that optimally combines pointwise model and observational errors. The JSBACH model error is unknown as is the (pointwise) observation error. We could use a general type of error estimate (such as that of 20 % of the flux value) for the observations but would have to include a minimal site and instrumentation-dependent precision. In this study, the full error is treated as Gaussian white noise. Because of these limitations, we are calling and defining our likelihood as a cost function. It is calculated with the same parameter values for each site, using site-specific daily measurements with the gap-filled, low-quality and winter (between the 315th and the 75th day of the year) values removed (resulting in N_{ET} and N_{GPP} points). These site-level estimates are averaged to produce the actual cost function, which is then returned for the algorithm to produce an estimate that is independent of the characteristics of any single site.

The cost function (Eq. 9) in our simulations is based on the normalised mean squared error (NMSE) estimates of the daily GPP and the daily ET. The residual of each variable is divided by the mean of observations, as has been previously done by, e.g., Mäkelä et al. (2016), Knauer et al. (2015), Groenendijk et al. (2010) and Trudinger et al. (2007). We make use of this approach since we needed to balance two series of different magnitudes (ET and GPP). The residuals are additionally divided by the (site-specific) number of observations so that the cost function is not biased towards any-specific site. The cost function (without the normalisation) can be interpreted as a negative log-likelihood function with a (Gaussian) error term equal to the observational mean.

$$\mathrm{cf}_1 = \overbrace{\frac{1}{N_{\mathrm{ET}}}\sum\left(\frac{\mathrm{ET_{mod}}-\mathrm{ET_{obs}}}{\overline{\mathrm{ET_{obs}}}}\right)^2}^{\mathrm{NMSE_{ET}}} + \overbrace{\frac{1}{N_{\mathrm{GPP}}}\sum\left(\frac{\mathrm{GPP_{mod}}-\mathrm{GPP_{obs}}}{\overline{\mathrm{GPP_{obs}}}}\right)^2}^{\mathrm{NMSE_{GPP}}} \quad (9)$$

We also use a modified version of this cost function, where the NMSEs are weighted by factors based on coefficients of determination (r^2) defined in Eq. (8). This latter cost function is only used during the separate drought period optimisation for Hyytiälä. During the drought we are more interested in the correct timing of the change in GPP and ET fluxes, rather than the size of the actual change. The aim is to correctly reproduce the changes in the water use efficiency (WUE) of plants, which we interpret here as the pointwise ratio of (ecosystem level) GPP to ET. The NMSE values ensure that the overall amplitude of the fluxes will remain satisfactory.

$$\mathrm{cf}_2 = (1-r^2_{\mathrm{ET}})\mathrm{NMSE_{ET}} + (1-r^2_{\mathrm{GPP}})\mathrm{NMSE_{GPP}} \quad (10)$$

3 Results

First we present the performance of the APIS algorithm and the parameters themselves, followed by site and stomatal conductance model-specific results and finally an examination of the Hyytiälä drought event in 2006. For simplicity, we use the name of the stomatal conductance model to refer to the JSBACH model utilising that stomatal conductance formulation.

The evolution of the APIS algorithmic process is presented in Fig. 1 for three parameters from the calibration of the Bethy model. The chosen parameters highlight different levels of identifiability for the algorithm (with the given cost function). The first parameter (f_{C3}) shows a well-identifiable situation, where the algorithm quickly locates the area of high probability. The second parameter (θ_{dr}) is also identifiable, but the speed of convergence is diminished. The last example (C_{decay}) represents situations where the parameter is not constrained. We have included images of the APIS chains for the other parameters as Supplement S1 along with parameter posterior estimates at 20 iterations with the Bethy and Ball–Berry formulations.

We also report the results of the Gelman–Rubin (Gelman and Rubin, 1992) and δ tests in Table 4. Both of these tests indicate that the algorithm is performing well at 20 iterations – the values of $\hat{R} \approx 1$, which means that further simulations are unlikely to improve the variance estimates. However, for some parameters, the convergence of the global estimate is slow (as also seen in the Supplement Fig. S1 for, e.g., τ, c_{b} and q). The APIS sampling process did not reveal any multimodal distributions and thus provided suitable initial conditions for the optimisation.

3.1 Optimised parameters

The results of the optimisation process are gathered in Table 5. There is an overall agreement on the values of the most prevalent parameters (see the bold and italic characters in Table 5 between the models). Most notably, the permanent wilting point (θ_{pwp}) and the point above which transpiration is unaffected by soil moisture stress (θ_{tsp}) have been significantly lowered. The LoGro-P parameters, which affect the timing of the spring and autumn events, are expected to contribute only little to the cost function. The coniferous evergreen trees do not shed all their leaves for winter, and therefore the timing of the bud burst is not as critical as for, e.g., deciduous trees. Additionally, because of the existing foliage, the state of acclimation parameter τ that depicts the reduction in carbon assimilation in the early spring likely dominates the phenology parameters that determine when new leaves start to grow.

Some of the parameters have converged to their limiting values, which can reflect deficiencies in the model structure or the preset parameter ranges. Convergence to the boundary can also be a problem in model calibration, but in this experiment, the algorithms were able to cope with the situation as APIS located the area of high probability and the optimiser located the maxima. The different parameter effectiveness levels reported in Table 5 can be roughly equated to the identifiability situations in Fig. 1. The effectiveness levels are highly situational (e.g. they depend on the sampling limits in Table 2 given for each parameter) and merely reflect the parameter identifiability in the APIS process. Low effectiveness complements the test results in Table 4, as the tests may indicate good performance for a parameter (e.g. for S_{range}) that is ineffective in the simulations.

3.2 Annual cycles

We present the average annual cycles for the validation period and for all sites in Fig. 2 using the Bethy formulation that is part of the standard model. The annual cycles generated with the other stomatal conductance models are added as Supplement S2. The parameters of the regression lines (b and r^2) between the measured and modelled ET and GPP fluxes of all the models are gathered in Table 6. These indicators have been calculated using all corresponding values regardless of the quality of the data. The sites are in the same order as in Table 1 with the six calibration sites first, followed by the four sites used only for validation. We have also included a supporting synthesis of the b and r^2 values between the model simulations with the default and optimised parameter values as Supplement S3.

The optimisation has improved the model bias and the correlation coefficients for the GPP in Fig. 2 for nearly every site, with the exception of deteriorating bias for Poker Flat (US-Prr) and Zotino (RU-Zot). Additionally, the improvement in the timing of the springtime increase in the GPP

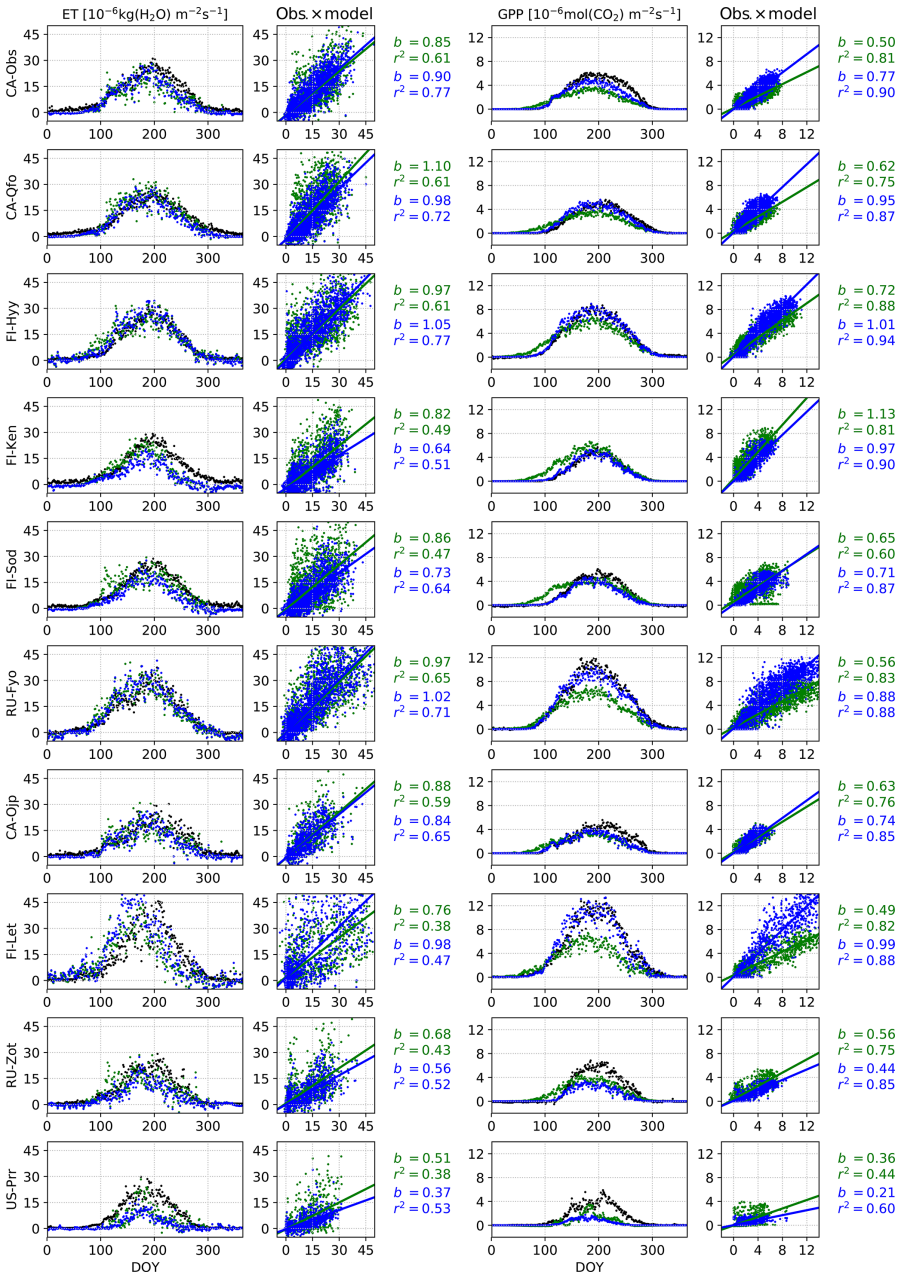

Figure 2. Validation period average annual cycles of evapotranspiration and gross primary production; observations (black) and the model using the Bethy stomatal conductance formulation with default (green) and optimised (blue) parameterisation. Also presented are daily model values cross plotted against observations with corresponding slope of the regression line (b) and the coefficient of determination (r^2).

Table 4. Parameter scale reduction $\hat{R}$ (at APIS iteration) and stability δ (threshold number of iterations) estimates from the Bethy simulations.

	$V_{C,max}$	α	τ	c_b	f_{C3}	q	θ_{dr}	θ_{hum}	θ_{pwp}
$\hat{R}$ at 20	1.12	0.99	1.02	0.99	1.0	0.99	1.0	1.3	1.08
$\hat{R}$ at 100	1.3	1.03	1.25	1.16	1.03	1.08	1.03	1.52	1.16
δ ($\pm$0.05)	20	21	27	40	0	36	18	14	17
	θ_{tsp}	p_{int}	s_m	w_{skin}	C_{decay}	S_{min}	S_{range}	T_{alt}	T_{ps}
$\hat{R}$ at 20	0.99	1.01	0.99	1.0	0.99	0.99	0.99	0.99	0.99
$\hat{R}$ at 100	1.06	1.13	1.0	0.99	0.99	1.0	0.99	0.99	0.99
δ ($\pm$0.05)	26	35	8	0	12	22	0	1	0

Table 5. Parameter default and optimised values for the calibration period with corresponding cost function value. The values written in boldface were the most effective and the italic values the least effective for the given experiment. Also presented are the fixed parameter values for the drought period optimisation, with "opt" referring to the use of the corresponding optimised value from this table.

Parameter	Def.	Base	Bethy	BB	Leu	F&K	USO	Dry set
$V_{C,max}$	62.5	48.4	57.1	55.4	49.7	50.8	**50.5**	52.0
α	0.28	0.318	0.318	0.319	0.317	0.319	0.318	0.318
τ	10.0	14.6	15.0	14.8	14.9	14.7	14.8	14.8
c_b	5.0	5.4	4.1	6.7	*4.4*	*4.3*	4.6	5.0
f_{C3}	0.87	**0.75**	**0.83**	–	–	–	–	Table 7
q	0.0	*0.03*	0.94	*0.62*	*0.60*	*0.82*	*0.65*	Table 7
g_0	1.0×10^{-3}	–	–	4.7×10^{-3}	4.7×10^{-3}	4.4×10^{-3}	4.2×10^{-3}	Table 7
g_1	Table 3	–	–	**9.9**	**8.8**	10.9	1.6	Table 7
a	2.8	–	–	–	–	3.2	–	opt
d	80	–	–	–	–	**71**	–	opt
θ_{dr}	0.9	0.86	*0.65*	0.88	0.83	0.8	*0.90*	0.85
θ_{hum}	0.5	**0.2**	**0.2**	**0.21**	**0.2**	**0.2**	**0.2**	Table 7
θ_{pwp}	0.35	**0.16**	**0.15**	**0.17**	**0.15**	**0.16**	**0.15**	Table 7
θ_{tsp}	0.75	**0.31**	**0.35**	**0.3**	**0.31**	**0.32**	**0.33**	Table 7
p_{int}	0.25	0.35	0.35	0.35	0.35	0.35	0.35	0.35
s_m	5.9×10^{-3}	0.099	*0.094*	*0.097*	*0.098*	0.097	*0.078*	0.097
w_{skin}	2.0×10^{-4}	$\mathbf{3.7 \times 10^{-4}}$	$\mathbf{3.1 \times 10^{-4}}$	$\mathbf{3.5 \times 10^{-4}}$	$\mathbf{3.6 \times 10^{-4}}$	$\mathbf{3.3 \times 10^{-4}}$	$\mathbf{3.2 \times 10^{-4}}$	3.4×10^{-4}
C_{decay}	13.0	*17.0*	*22.2*	*23.3*	*23.3*	*24.9*	*13.9*	opt
S_{min}	10.0	*29.2*	*26.3*	*10.7*	6.3	*26.1*	6.3	opt
S_{range}	150	*247*	*176*	*162*	*157*	*202*	*223*	opt
T_{alt}	4.0	2.0	2.8	5.8	8.2	2.5	8.3	opt
T_{ps}	10.0	*18.6*	24.4	3.8	3.2	15.0	3.1	opt
cf_1		0.571	0.531	0.521	0.529	0.518	0.528	

is apparent. All of the correlation coefficients for the ET in Fig. 2 have also been improved but the model bias has mostly increased.

3.3 Drought event

The resulting parameter values, from the optimisation during the drought conditions in Hyytiälä (FI-Hyy) in the summer of 2006, are presented in Table 7. Setting the maximum carboxylation rate to a constant value ($V_{C,max} = 52.0$) enabled the full use of the dynamical range of q – the idea was to ensure that $V_{C,max}$ does not dominate the optimisation, any value for q is possible, and it is able to influence the outcome. The LoGro-P parameters and τ were fixed to their optimised values, presented in Table 5, as they should not be affected by the drought. Likewise, the values of other parameters (not presented in Table 7) were set as compromises between the stomatal conductance formulations.

We can now compare the parameter values in Table 7 to those in Table 5. The values of the relative humidity parameter (θ_{hum}) and the residual stomatal conductance (g_0) have remained nearly unchanged, but the rest of the parameters have quite varied values. The leaf internal-to-external CO_2 concentration (f_{C3}) as well as the slope of the stomatal conductance (g_1) are at the lower bound (expect g_1 for BB – Ball–

Table 6. Slope of the regression line (b) and the coefficient of determination (r^2) for the different stomatal conductance formulations during the validation period with the optimised parameters. We have written the best values of b and r^2 in boldface for each site and italicised the abbreviations of the separate validation sites.

	Evapotranspiration (ET)											
	b						r^2					
Site	Base	Bethy	BB	Leu	F&K	USO	Base	Bethy	BB	Leu	F&K	USO
CA-Obs	0.91	0.9	**0.91**	0.86	0.81	0.76	0.75	**0.77**	0.76	0.76	0.75	0.74
CA-Qfo	0.96	0.98	**0.99**	0.92	0.89	0.83	0.71	**0.72**	0.7	0.71	0.7	0.69
FI-Hyy	0.97	1.05	1.07	0.95	**0.98**	0.79	0.73	**0.77**	0.77	0.75	0.77	0.69
FI-Ken	0.54	**0.64**	0.62	0.56	0.58	0.48	0.48	0.51	**0.52**	0.49	0.51	0.45
FI-Sod	0.64	0.73	**0.74**	0.63	0.64	0.56	0.58	**0.64**	0.61	0.6	0.62	0.55
RU-Fyo	0.98	1.02	1.01	0.98	**0.99**	0.85	0.7	0.71	0.71	0.71	**0.71**	0.7
CA-Ojp	0.8	**0.84**	0.84	0.75	0.72	0.67	0.64	**0.65**	0.64	0.65	0.64	0.63
FI-Let	1.09	0.98	1.08	1.04	**1.01**	0.94	0.49	0.47	0.49	0.5	**0.51**	0.48
RU-Zot	0.49	0.56	**0.56**	0.47	0.46	0.41	0.45	**0.52**	0.5	0.47	0.48	0.41
US-Prr	0.38	0.37	**0.42**	0.35	0.33	0.35	0.48	0.53	**0.53**	0.46	0.44	0.43
Best values	0	2	5	0	3	0	0	6	2	0	2	0
	Gross primary production (GPP)											
	b						r^2					
Site	Base	Bethy	BB	Leu	F&K	USO	Base	Bethy	BB	Leu	F&K	USO
CA-Obs	**0.83**	0.77	0.82	0.81	0.81	0.77	0.87	0.9	0.89	0.89	**0.91**	0.9
CA-Qfo	0.97	0.95	**0.98**	0.96	0.96	0.9	0.84	0.87	0.85	0.86	**0.88**	0.87
FI-Hyy	1.02	**1.01**	1.05	1.03	1.06	0.98	0.94	0.94	0.94	0.95	**0.95**	0.95
FI-Ken	0.9	0.97	**0.97**	0.93	0.95	0.9	0.93	0.9	0.9	0.93	0.93	**0.94**
FI-Sod	0.66	0.71	**0.71**	0.67	0.69	0.65	0.88	0.87	0.86	0.89	**0.9**	0.9
RU-Fyo	0.95	0.88	0.91	0.96	**0.98**	0.91	0.89	0.88	0.88	0.91	0.91	**0.91**
CA-Ojp	0.72	0.74	**0.75**	0.7	0.69	0.66	0.83	0.85	0.84	0.85	0.86	**0.86**
FI-Let	1.27	**0.99**	1.09	1.25	1.26	1.21	0.93	0.88	0.89	0.94	0.94	**0.94**
RU-Zot	0.42	**0.44**	0.44	0.42	0.42	0.4	0.86	0.85	0.84	0.88	**0.88**	0.88
US-Prr	0.2	**0.21**	0.21	0.2	0.19	0.19	0.62	0.6	0.6	0.62	**0.63**	0.62
Best values	1	4	4	0	1	0	0	0	0	0	6	4

Berry). Noticeably, the unified stomatal optimisation model (USO) only changes the values of θ_{tsp} and q and leaves the rest of the parameters almost untouched.

The changes these different parameterisations have on the model output are visualised in Fig. 3. All of the stomatal conductance models, with default parameterisation, suffer from ET values that are too low before (and during) the actual drought. This behaviour was corrected during the general optimisation but has partially re-emerged with the dry-period parameters for the Baseline, Ball–Berry, Leuning, and to a lesser degree the Friend and Kiang formulations. Most of the models also exhibit ET values that are too high during the actual drought with the generally optimised parameter values. This behaviour was also corrected with the dry-period optimisation, but the Baseline and especially the Bethy model now suffer from too strong a drawdown of ET. These models also demonstrate the drawdown that is too strong for the GPP. The GPP itself was greatly improved with both optimisations and for all models. The dry-period optimisation of the USO also managed to correct the erroneous GPP of the general optimisation during the actual drought, whereas the GPP of other formulations has remained roughly the same as with the general optimisation. The USO formulation results in the best fits for r^2 and b with the dry-period optimisation.

The Bethy and the USO models demonstrate the most variability in the β-function values in Fig. 3 (rightmost panels), for the dry-period optimisation. We selected these two stomatal conductance formulations to examine the changes to the WUE of plants during the extended dry period. The highlighted observations in Fig. 4c and f show a clear path of development for the drought where the observations imitate the letter δ. The colourings follow the β-function values in Fig. 3 between the red vertical lines. Both observational colourings (same as the model colouring) are similar and depict, ini-

tially, a linear decrease in both ET and GPP, followed by a rapid decline in ET and a delayed decline in GPP. The recovery of plants from the drought can also be seen as the colouring starts to turn lighter. The models depict a more linear response of GPP to ET as the drought develops, although with USO we can see a few more similarities in the pattern of the values.

Finally, we used both optimised parameter sets (Tables 5 and 7) to produce the ET and GPP cycles for all sites and stomatal conductance models. This analysis (not shown) verified that in general conditions, the Table 5 parameter values produced better estimates in general. The b and r^2 values for the ET were systematically better for all stomatal conductance formulations (except one). There was some variation in the indicators for the GPP, where approximately a third of the values (of mostly r^2) are better with the dry-period parameter set. These differences are mostly attributed to increased model bias (decreased b) that is explained by the lower values of g_1. Overall, the more general optimisation provided systematically better or comparable results to the dry-period optimisation. The exception is the USO formulation, which had an approximately 1 : 1 distribution of best values for both variables in between the parameter sets.

4 Discussion

We will first discuss the validity of our approach and the simulation setup, followed by an examination of the success of the modifications made to the model, and close with some further remarks on the parameter values.

4.1 Validity of the simulations

Before we calibrated the model, we fixed the limiting value for LAI and adjusted the site-specific vegetative area fractions to reproduce the measured site-level maximum of LAI. In the simulations, we focused on boreal coniferous forests, where light penetration is deep and the light conditions are homogenous – consequently we could assume a homogenous leaf distribution. Furthermore, the JSBACH model takes into account leaf clumping, and we can assume the leaf orientation and shape to be similar throughout the study sites. Therefore, we argue that reproducing the site-level maximum of LAI is appropriate approach in this study. Together with parameter calibration, it has resulted in improved ET and GPP fluxes as can be verified from the b and r^2 values in Fig. 2. The improvements in b and r^2 are mostly seen in the GPP flux, which can be explained by the fact that the stomatal conductance in JSBACH is primarily resolved for carbon assimilation, and the same conductance is then used for transpiration (Eq. A8). Additionally, GPP is derived from the EC measurements by flux partitioning – this tends to remove some of the flux instabilities (that are still present in the ET).

We encountered difficulties in reproducing the fluxes for the validation sites with low LAI (i.e. RU-Zot and US-Prr). This can be a consequence of the area scaling as the adjustment linearly changes the proportions between vegetative area and bare soil. Another reason is the lack of the site understorey in these simulations. For example, approximately half of the CO_2 fluxes (and consequently roughly half of the GPP) for Poker Flat are produced by the site understorey (Ikawa et al., 2015). Additionally, there are also many parameters describing site-specific soil properties (such as porosity) that were not part of the optimisation and may be inaccurate. These effects may also be pronounced due to the changes in parameters affecting soil moisture as well as the area scaling.

There were no clear differences between sites dominated by pine or spruce. Neither did we notice any particular effect on the bias, NMSE or correlation coefficient that could be explained by geographical location, stand age,s or annual precipitation or temperature. We optimised the model for individual (calibration) sites as well (not shown). Mostly this changed the values of parameters (such as $V_{\mathrm{C,max}}$ and g_1), affecting the amplitude of the modelled fluxes. These parameters can be viewed to be more site-specific, a characteristics that is reduced in a multi-site calibration – the possibility of highly site-specific properties (and parameter values) can also explain the difficulties in reproducing the validation site observations. We are omitting these results as single-site optimisation can be viewed as overfitting the model and the results do not provide any additional insights.

The APIS performance tests (Gelman–Rubin and δ) indicate that the algorithm is performing well at 20 iterations, but the convergence of the global estimate for some parameters is slow. This is mostly a direct result of the normalisation of the cost function that inflates the target distribution, which reduces the parameter sensitivity to observations and gives too much weight to the initial locations and draws. Without the normalisation, the algorithm would also converge faster. Additionally, APIS is meant to examine the full target distribution with only some sequentiality – 20 iterations (or less) should be sufficient for APIS to locate the modes of the target. In longer APIS simulations, the global estimate would likely benefit from discarding the first half of the samples, but this would require the estimate to be recalculated at each iteration (from the drawn samples) as it could not be calculated iteratively.

4.2 Delayed effect of temperature

We modified the JSBACH model by introducing the delayed effect of temperature for photosynthesis to restrain the respiration and photosynthesis of conifers in spring. The effect of this (delayed increase in GPP) is apparent in the annual GPP cycles of CA-Qfo, FI-Hyy, FI-Ken, FI-Sod and RU-Zot in Fig. 2. The delay is in place for the other sites as well, but the effect is less apparent in the figure. This delay is to a

Table 7. Optimised parameter and corresponding cost function values with different stomatal conductance formulations for the extended dry period.

Parameter	Def.	Base	Bethy	BB	Leu	F&K	USO
f_{C3}	0.87	0.7	0.7	–	–	–	–
q	0.0	0.09	0.0	0.15	0.57	0.16	0.30
θ_{tsp}	0.75	0.57	0.46	0.48	0.44	0.45	0.41
θ_{pwp}	0.35	0.40	0.38	0.27	0.23	0.28	0.16
θ_{hum}	0.5	0.2	0.2	0.2	0.2	0.2	0.2
g_0	Table 3	–	–	4.9×10^{-3}	5.0×10^{-3}	3.8×10^{-3}	4.6×10^{-3}
g_1	Table 3	–	–	7.5	6.0	7.0	1.5
cf_2		0.42	0.44	0.39	0.41	0.41	0.41

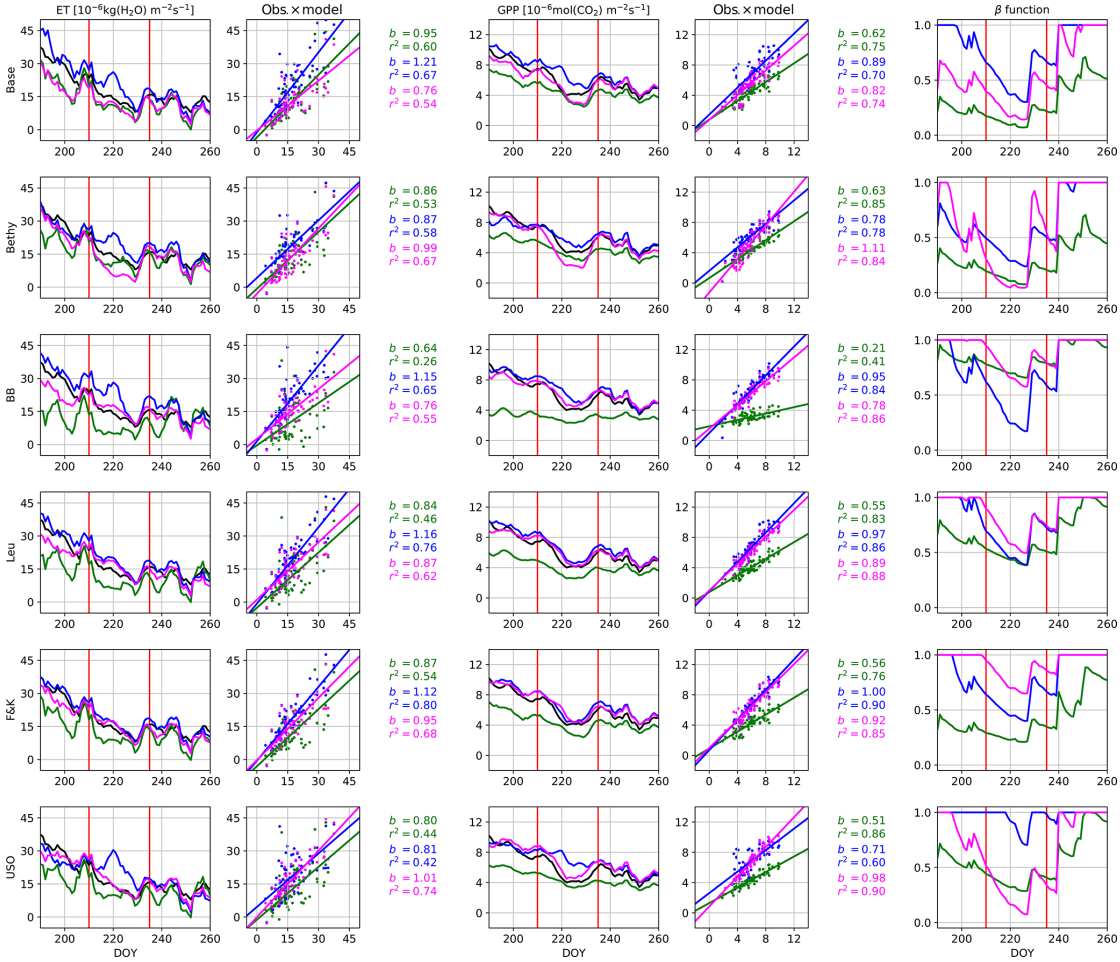

Figure 3. Hyytiälä site drought in summer 2006. The time series for evapotranspiration and gross primary production are 5 d running averages and for β-function daily values. The observations are plotted in black and the model with default parameterisation in green, calibration period optimisation in blue and the dry-year optimisation in magenta. The red vertical lines indicate the start and end of the actual drought.

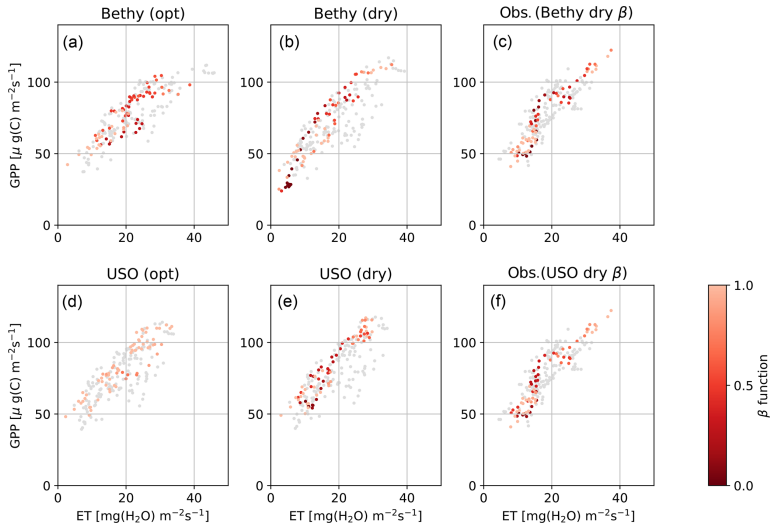

Figure 4. Hyytiälä site water use efficiency for the Bethy and USO formulations. Scatter-plotted are the dry-period 5 d running averages of ET and GPP, coloured by the intensity of the drought (β function). Panels **(a, d)** depict the model with the more generally optimised parameter values and **(b, e)** with the drought optimisation, and **(c, f)** present the corresponding observations, coloured by the same intensity as in **(b, e)**. The grey points are from the corresponding time during the 2 previous years.

lesser extent also reflected in transpiration, and consequently in ET, as can be seen, e.g., at FI-Hyy and FI-Sod – for other sites this effect is not clear. The correction in the ET values can lead to an increase in model bias as is the case with Sodankylä (FI-Sod), where the autumn values that are too low in the default model were previously compensated for by springtime values that too high (in the sense of annual ET). Correcting the springtime behaviour leads to an increase in bias, but this should not be viewed as a fault in the optimisation as the model was previously mitigating an erroneous behaviour (too low an autumn ET) with another (too high a springtime ET).

Mäkelä et al. (2004) used a linear dependency of photosynthetic efficiency to the state of acclimation and reported 13.75 d to be the best fit for the adjustment period length (τ). Kolari et al. (2007) utilised a sigmoidal relation and reported the value of 8 d but noted that the range of values resulting in a good fit was large (5–10.4 d). Linkosalo et al. (2014) came to a similar conclusion when they encountered a near-flat distribution for τ in the range of 1–12 d. In our simulations τ exhibits larger optimal values (nearly 15 d), which is most likely due to the model adapting to the multi-site calibration (as sites have different characteristics, a longer acclimation period accounts better for these variations).

4.3 Stomatal conductance models

We examined the model behaviour with six stomatal conductance formulations, and the resulting b and r^2 values are presented in Table 6. The best performance (bolded values) in simulated ET is achieved by the BB model for bias and the Bethy formulation for r^2. These two models also share the best performance in the GPP bias, whereas the best r^2 values for the GPP are demonstrated by the F&K (Friend and Kiang) model, followed by the USO formulation. Calculating the number of best values demonstrated by each model, we obtain that the best performance is shared by the Bethy (12) and F&K (12) formulations, followed by the BB (11) model. However, we note that some of the "best values" are only marginally better than comparable values. Additionally, we used two more parameters (a and d) for the F&K formulation than for the other Ball–Berry formulations. Likewise, we could have, for example, included the factor D_0 (Eq. B3) in the optimisation, which would have likely improved the performance of the Leuning model. Similarly to the results by Knauer et al. (2015), based on this (general) calibration, there is no clear single candidate for the best stomatal conductance formulation.

The model behaviour was also examined during the Hyytiälä drought of 2006. Some of the parameter values were kept fixed during these simulations; most of the fixed parameters should not affect the drought period calibration, but there are exceptions, such as the maximum carboxylation rate $V_{\mathrm{C,max}}$. It can be argued that, e.g., both the parameters $V_{\mathrm{C,max}}$ and g_1 should decrease (Egea et al., 2011) during the drought, but we decided to fix $V_{\mathrm{C,max}}$ to get a better response for q. The best fit to the observations was achieved by the USO formulation, as seen in Fig. 3, with remarkably similar

parameter values to the general optimisation. The USO was also able to (somewhat) replicate the "δ" shape of the drought in Fig. 4.

The stomatal conductance function ($g_s = g_0 + c\beta g_1$) also incorporates the soil water parameters θ_{tsp} and θ_{pwp} in the form of the β function as portrayed in Eq. (2). The changes in the values of these parameters (mostly g_1, θ_{tsp} and θ_{pwp}) are intertwined. During the drought, the decrease in the optimised values of g_1 is expected as the plants close their stomata to minimise the loss of water by transpiration (Egea et al., 2011; Zhou et al., 2013). The same effect is also achieved by increasing the values of θ_{tsp} and θ_{pwp} as this decreases the values of the β function. The higher values of g_1 during the more general optimisation are better reflected by Franks et al. (2018), whereas the lower values during the drought are more in accordance with physiological observations by Egea et al. (2011). Likewise, Lin et al. (2015) found higher values for g_1 (both boreal area and gymnosperm trees) using the USO.

In general, the site-level estimates of (g_0 and) g_1 are sensitive not only to the stomatal conductance formulation but also, e.g., to the structure of the underlying model and the value of other parameters, such as maximum carboxylation rate ($V_{C,max}$). (Wang, 1996, Table 1, Control) reported $g_1 = 3.78$, using a Leuning model similar to ours, where $(1 + D_S/D_0)$ is replaced by D_S. Thum et al. (2007) approximated g_1^{BB} to be 5 for Sodankylä while estimating the variation in the values of $V_{C,max}$ and maximum rate of electron transport J_{max}. We would suggest that the limiting values θ_{pwp} and θ_{tsp} should be optimised or fixed before introducing additional tuning factors such as mesophyll conductance or scaling the β in multiple ways in the stomatal conductance formulations (Egea et al., 2011). Our simulation setup for q corresponds to the configuration 5 (C5) by Egea et al. (2011), with variables $q = q_B$ and fixed value $q_S = 1$.

4.4 Parameter values

Some of the parameters in this study have been calibrated before by, e.g., Kattge et al. (2009) and Knorr et al. (2010). Our approach differs from these as we required the model to reproduce the site-level maximum of LAI. In contrast, e.g., Knorr et al. (2010) found the structural limit for (all-sided) LAI to be 4.2, which is considerably lower than the measured LAI for many of the sites in Table 1. Our approach directly scales the vegetative area, so it also scales GPP and also the amount of rain available for plants (as rain is directed to bare soil and vegetative area). This means that the parameter values should not be directly compared without taking the different paradigms into account. However, our optimised $V_{C,max}$ values are in between 62.5 reported by Kattge et al. (2009) and 29.3 by Knorr et al. (2010) and are in line with the yearly cycle presented by Ueyama et al. (2016).

The exponential scaling factor q in Eq. (A1) of the β function (Eq. 2) was revealed to be ineffective in our optimisation as indicated in Table 5. In our simulations, this situation arises as the effective range of the β function has been lowered by reducing θ_{pwp} and θ_{tsp}. The actual soil moisture is rarely below the fraction θ_{tsp}, so q is constrained with a very limited number of datapoints and thus only has a minimal effect on the fluxes and the cost function. Therefore, the values presented for q in Table 5 can be unreliable and even unrealistic. This situation is remedied in the drought period optimisation when the soil moisture is low. The resulting values for q in Table 7 have a wide dispersion, although they are mostly at the lower end. This signifies that the additional GPP reduction is mostly gradual, with a steep decrease near the permanent wilting point θ_{pwp}.

The values of soil water parameters are closely grouped in the optimisations except for the values of θ_{pwp} during the drought. This can occur due to a larger impact of the different stomatal conductance formulations on the accumulating soil water content than assumed – this can also be seen from the differences in the β-function values in Fig. 3. Furthermore, the values of θ_{tsp} and θ_{pwp} have been considerably lowered from their default values in both optimisations. This change can be perceived in at least two different ways. Either the boreal forests are not generally limited by soil moisture stress (except in the case of extreme drought) or the water retention capabilities of the soil (in the model) have been systematically overestimated. The latter seems unlikely, in the light of results by, e.g., Gao et al. (2016).

5 Conclusions

APIS is a recent method, capable of estimating complicated multidimensional probability distributions using a population of different proposal densities. The algorithm was able to produce reasonably stable estimates for most parameters quickly. Prior to calibrating the model, we adjusted the site-specific vegetative area fractions to reproduce the measured site-level maximum of LAI. This practical approach resulted in improved ET and GPP fluxes, although we encountered difficulties in replicating these for sites with low LAI. The model parameters were optimised simultaneously for all sites without any additional site-level tuning. The parameters that were most effective in the optimisation processes were consistent for all stomatal conductance formulations.

The introduction of the S function, to delay the start of the vegetation active season, has corrected the springtime increase in GPP for conifers throughout the sites used in this study. The parameters θ_{tsp} and θ_{pwp}, which set the range for the soil moisture stress function β, were both systematically lowered and optimised to nearly identical values for all stomatal conductance models. The low effective range for the β function rendered the experimental parameter q nearly ineffective in the more general optimisation. The dry-period optimisation increased the effective range of the β function and the importance of q, which resulted in a highly nonlinear (additional) reduction in the net assimilation rate. Overall, this

fact and both optimisations indicate that boreal forest transpiration is not limited by soil moisture stress under normal conditions.

The optimisation improved the predictive skill of the model with all stomatal conductance formulations as was seen during the validation period. The Bethy, Ball–Berry, and Friend and Kiang versions were the most in agreement with the observations, although the differences between these and the other formulations were small. Most of the model versions had some problems during the extended dry period, and the best b and r^2 values were achieved by the unified stomatal optimisation model. Additionally, the optimised parameter values of the USO for the dry period were the most similar (of all stomatal conductance formulations) to those of the more general optimisation.

Code and data availability. The data required to calibrate and validate the model are originally part of the FLUXNET2015 dataset that can be accessed through the FLUXNET database (https://doi.org/10.17616/R36K9X, Baldocchi et al., 2011). Our modified dataset, containing the forcing data and the observations used in this article, is available through Zenodo portal (https://doi.org/10.5281/zenodo.3240954, Mäkelä, 2019). The data depicting the simulations (parameter draws, cost function values, etc.) have been added as a Supplement. The JSBACH model (branch: cosmos-landveg-tk-topmodel-peat, revision: 7384) can be obtained from the Max Planck Institute for Meteorology, where it is available for the scientific community under the MPI-M Sofware License Agreement (http://www.mpimet.mpg.de/en/science/models/license/, last access: 16 September 2019). The modifications to the model, described in this paper, have been uploaded to Github, and they can be accessed by contacting the authors at jarmo.makela@fmi.fi (after access to the actual model has been approved). For any questions, we encourage you to contact the authors at jarmo.makela@fmi.fi.

Appendix A: Parametric equations within JSBACH

In this Appendix we present the most relevant equations that are governed by the parameters in Table 2. The Appendix is divided into sections that coincide with the parameter groups.

A1 Photosynthesis

The Farquhar model (Farquhar et al., 1980) is based on the observation that the assimilation rate in the chloroplast is limited either by the carboxylation rate (V_C), induced by the Rubisco enzyme, or the light-limited assimilation rate (J_E). The total rate of carbon fixation is reduced by the amount of dark respiration (R_d), resulting in net assimilation rate (A_n). The experimental scaling factor β^q (Egea et al., 2011) is based on soil moisture stress in Eq. (2), which takes effect ($\beta < 1$) when soil moisture is significantly reduced. This scaling is used by all stomatal conductance formulations. We have also introduced here in equation form the actual reduction to photosynthesis by γ from the delay in the start of the vegetation active season in Eq. (1).

$$A_n = \beta^q(\min(\gamma V_C, J_E) - \gamma R_d) \tag{A1}$$

Oxygenation of the Rubisco molecule reduces the carboxylation rate, which is given as

$$V_C = V_{C,max}\frac{C_i - \Gamma_\star}{C_i + K_C(1 + O_i/K_O)}. \tag{A2}$$

Here C_i and O_i are the leaf-internal CO_2 and O_2 concentrations, $\Gamma_\star$ is the photorespiratory CO_2 compensation point, and K_C and K_O are Michaelis–Menten constants parameterising the dependence on CO_2 and O_2 concentrations. Furthermore, leaf-internal CO_2 concentration depends on the external (ambient) concentration C_a (in the Baseline and Bethy formulations and unstressed conditions) by

$$C_i = f_{C3}C_a. \tag{A3}$$

Likewise, the light-limited assimilation rate can be expressed as a function on electron transport rate (J), which is a function of radiation intensity (I) in the photosynthetically active band, the maximum electron transport rate (J_{max}) and the quantum efficiency for photon capture (α):

$$J_E = J(I)\frac{C_i - \Gamma_\star}{4(C_i + 2\Gamma_\star)}, J(I) = J_{max}\frac{\alpha I}{\sqrt{J_{max}^2 + \alpha^2 I^2}}. \tag{A4}$$

A2 Soil water

In JSBACH the soil water budget is based on several reservoirs (e.g. skin, soil, bare soil, rain intercepted by canopy), and the different formulations are plentiful. We present here only the most crucial of these. Changes in volumetric soil water (θ_s, not to be confused with relative soil water content $\theta = \frac{\theta_s}{\theta_{fc}}$) due to rainfall ($R$), evapotranspiration (ET), snow melt (M), surface runoff (R_s) and drainage (D), are calculated with a geographically varying maximum field capacity (θ_{fc}) and soil water density (ρ_w).

$$\rho_w\frac{\partial\theta_s}{\partial t} = (1 - p_{int})R + ET + M - R_s - D \tag{A5}$$

The interception parameter (p_{int}) also affects the amount of water intercepted by vegetation and bare soil, which further affects evaporation and transpiration. The skin reservoir is limited by w_{skin}, and excess water is transferred to soil water. Likewise when the soil water content (θ) is greater than parameter θ_{dr}, the excess water is rapidly drained (in addition to the limited drainage below this threshold), where d, d_{min} and d_{max} are constant parameters:

$$D = d_{min}\theta + (d_{max} - d_{min})\left(\frac{\theta - \theta_{dr}}{1 - \theta_{dr}}\right)^d, \theta \geq \theta_{dr}. \tag{A6}$$

Evaporation from wet surfaces (E_{ws}) depends on air density (ρ), specific humidity (q_a), saturation-specific humidity (q_s) at surface temperature (T_s) and pressure (p_s), and aerodynamic resistance (R_a). The aerodynamic resistance depends on heat transfer coefficient (C_h) and horizontal velocity (v_h).

$$E_{ws} = \rho\frac{q_a - q_s(T_s, p_s)}{R_a}, R_a = C_h|v_h|^{-1} \tag{A7}$$

Transpiration from vegetation (T_v) is likewise formulated but additionally depends on the stomatal resistance of the canopy (R_c), which is an inverse of the stomatal conductance, and as such, depends on which conductance model is used.

$$T_v = \rho\frac{q_a - q_s(T_s, p_s)}{R_a + R_c} \tag{A8}$$

Evaporation from dry bare soil (E_s) also has an added dependence on surface relative humidity (h_s) calculated from soil dryness:

$$E_s = \rho\frac{q_a - h_s q_s(T_s, p_s)}{R_a},$$
$$h_s = \max\left[\theta_{hum}(1 - \cos(\pi\theta)), \min\left(1, \frac{q_a}{q_s(T_s, p_s)}\right)\right]. \tag{A9}$$

The total evapotranspiration is a weighted average of E_{ws}, T_v and E_s, where the weights are based on fill levels of reservoirs and the vegetative fraction of the grid cell.

A3 Logistic Growth Phenology model

The parameters from the LoGro-P are mainly used to determine the spring and autumn events for JSBACH. To determine the date of the spring event, we first introduce a few additional variables, namely the heatsum $S_T(d)$, the number

of chill days $C(d)$ and the critical heatsum $S_{\mathrm{crit}}(d)$. $T(d)$ denotes the mean temperature at day d.

$$S_T(d) = \sum_{d'=d_0}^{d} \max(T(d') - T_{\mathrm{alt}}, 0) \tag{A10}$$

Heatsum $S_T(d)$ cumulates the amount of "heat" above the parameter T_{alt} after the previous growing season. The actual starting date d_0 of the summation need not be known since it is enough to start the summation "reasonably late" after the last growth season.

$$C(d) = \sum_{d'=d_{\mathrm{a}}}^{d} H\left(T_{\mathrm{alt}} - T(d)\right) \tag{A11}$$

The number of chill days is calculated as the number of days when the mean temperature is below T_{alt}. Here $H()$ denotes the Heaviside step function, and the summation starts at the day (d_{a}) of the last autumn event.

$$S_{\mathrm{crit}}(d) = S_{\mathrm{min}} + S_{\mathrm{range}} e^{-C(d)/C_{\mathrm{decay}}} \tag{A12}$$

The critical heatsum (S_{crit}) decreases as the number of chill days $C(d)$ increases, with an exponential memory loss parameter C_{decay}. The spring event happens when

$$S_T(d) \geq S_{\mathrm{crit}}(d). \tag{A13}$$

The autumn event requires the definition of one more variable, the (pseudo) soil temperature ($T_{\mathrm{s}}(t)$), which at time t is calculated as an average air temperature (T) with an exponential memory loss (T_{ps}). The autumn event occurs when T_{s} falls below a certain threshold. In the equation N is the normalisation constant and τ is the length of a time step.

$$T_{\mathrm{s}}(t) = \frac{1}{N} \sum_{n=-\infty}^{t} T(n) e^{-(t-n)\frac{\tau}{T_{\mathrm{ps}}}} \tag{A14}$$

Appendix B: Stomatal conductance formulations

In this Appendix we present the stomatal conductance model formulations used in this study. In the original JSBACH formulation, the Baseline model (Knorr, 1997), the photosynthetic rate is resolved in two steps. First the stomatal conductance under conditions with no water stress is assumed to be controlled by photosynthetic activity (Schulze et al., 1994). Here the leaf-internal CO_2 concentration is assumed to be a constant fraction ($C_{\mathrm{i,pot}} = f_{\mathrm{C3}} C_{\mathrm{a}}$) of ambient CO_2 concentration (C_{a}). This allows for an explicit resolution of the photosynthesis (Knorr, 1997). Then the impact of soil water availability is accounted for by a soil-moisture-dependent multiplier (β) that is identical for each canopy layer (Knorr, 1997).

$$g_{\mathrm{s,pot}} = \frac{1.6 A_{\mathrm{n,pot}}}{C_{\mathrm{a}} - C_{\mathrm{i,pot}}} \Rightarrow g_{\mathrm{s}} = \beta g_{\mathrm{s,pot}} \tag{B1}$$

After accounting for soil water stress, the net assimilation rate (A_{n}) and intercellular CO_2 concentration are (C_{i}) are recalculated using g_{s} and integrated over the leaf area index to produce canopy level estimates.

In the Bethy approach (Knorr, 2000), the unstressed canopy conductance ($G_{\mathrm{c,pot}}$) is calculated similarly to the Baseline model but is potentially further limited by the water supply function of the maximum transpiration rate ($T_{\mathrm{supply}} = \beta T_{\mathrm{max}}$). T_{max} is a fixed and predefined upper limit for transpiration as in Knauer et al. (2015).

$$G_{\mathrm{c}} = \begin{cases} G_{\mathrm{c,pot}} \frac{T_{\mathrm{supply}}}{T_{\mathrm{pot}}}, & T_{\mathrm{pot}} \geq T_{\mathrm{supply}} \geq 0 \\ G_{\mathrm{c,pot}}, & T_{\mathrm{pot}} < T_{\mathrm{supply}} \end{cases}$$

$$T_{\mathrm{pot}} = \rho \frac{q_{\mathrm{s}} - q_{\mathrm{a}}}{1/G_{\mathrm{a}} + 1/G_{c,\mathrm{pot}}} \tag{B2}$$

The potential (unstressed) transpiration rate (T_{pot}) is a function of air density (ρ), saturation-specific humidity (q_{s}) at given temperature and pressure, specific humidity (q_{a}), aerodynamic conductance (G_{a}), and unstressed canopy conductance ($G_{\mathrm{c,pot}}$). After this scaling, the net assimilation rate and intercellular CO_2 concentration are recalculated as in the Baseline model.

The Ball–Berry variants relate the stomatal conductance (g_{s}) to empirically fitted parameters g_0 (mol m^{-2} s^{-1}) and g_1 (unitless, except for g_1^{USO}, which has units of $\sqrt{\mathrm{kPa}}$) that respectively represent the residual stomatal conductance and the slope of the function. The stomatal conductance is a function of the net assimilation rate (A_{n}), the water stress factor (β) and the atmospheric CO_2 concentration (C_{a}). The original Ball–Berry formulation (Ball et al., 1987) also depends on relative humidity at leaf surface (h_{s}). In the Leuning model (Leuning, 1995), the CO_2 concentration is reduced by the CO_2 compensation point (Γ) as well as scaled by the vapour pressure deficit (D_{s}) and a constant (D_0) depicting the stomatal sensitivity to changes in D_{s}. The Friend and Kiang model (Friend and Kiang, 2005) adds an exponential dependency on the difference of specific (q_{a}) and saturation-specific humidity (q_{sat}) with empirically fitted constants $a = 2.8$ and $b = 80$. The unified stomatal optimisation model (Medlyn et al., 2011) also adds a dependency to the vapour pressure deficit (D_{s}).

$$\begin{aligned}
g_s^{\mathrm{BB}} &= g_0^{\mathrm{BB}} + g_1^{\mathrm{BB}} \beta \frac{A_{\mathrm{n}} h_{\mathrm{s}}}{C_{\mathrm{a}}} \\
g_{\mathrm{s}}^{\mathrm{Leu}} &= g_0^{\mathrm{Leu}} + g_1^{\mathrm{Leu}} \beta \frac{A_{\mathrm{n}}}{(C_{\mathrm{a}} - \Gamma)(1 + D_{\mathrm{s}}/D_0)} \\
g_{\mathrm{s}}^{F\&K} &= g_0^{F\&K} + g_1^{F\&K} \beta \frac{A_{\mathrm{n}} a^{-d(q_{\mathrm{sat}} - q_{\mathrm{a}})}}{C_{\mathrm{a}}} \\
g_{\mathrm{s}}^{\mathrm{USO}} &= g_0^{\mathrm{USO}} + 1.6 \left(1 + \frac{g_1^{\mathrm{USO}} \beta}{\sqrt{D_{\mathrm{s}}}}\right) \frac{A_{\mathrm{n}}}{C_{\mathrm{a}}}
\end{aligned} \tag{B3}$$

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/gmd-12-4075-2019-supplement.

Author contributions. The experiments were planned by JM, TA, TM and TT. JM ran the simulations and prepared the paper with contributions from co-authors. JK originally implemented the Ball–Berry type stomatal conductance formulations into JSBACH under SZ's supervision. JS maintained the framework for testing the algorithm. MA, AB, MH, AL, IM, HM and HK provided the site-level observations required in this study. TA, TM and TV extensively commented and revised the document.

Competing interests. The authors declare that they have no conflicts of interest.

Acknowledgements. This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia and USCCC. The FLUXNET eddy covariance data processing and harmonisation was carried out by the ICOS Ecosystem Thematic Center, AmeriFlux Management Project and Fluxdata project of FLUXNET, with the support of CDIAC, and the OzFlux, ChinaFlux and AsiaFlux offices.

Financial support. This research has been supported by the Jenny ja Antti Wihurin Rahasto, the NordForsk (grant no. 57001), the Academy of Finland (grant nos. 295874 and 307331), the ICOS-Finland (grant no. 281255) and the EU-Life+ (grant no. LIFE12 ENV/FI000409).

Review statement. This paper was edited by Hisashi Sato and reviewed by three anonymous referees.

References

Aurela, M., Lohila, A., Tuovinen, J., Hatakka, J., Penttilä, T., and Laurila, T.: Carbon dioxide and energy flux measurements in four northern-boreal ecosystems at Pallas, Boreal Environ. Res., 20, 455–473, 2015.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, Ch., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Leek, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw U, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities, B. Am. Meteorol. Soc., 82, 2415–2434, https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2, 2001.

Ball, J., Woodrow, I., and Berry, J.: A Model Predicting Stomatal Conductance and its Contribution to the Control of Photosynthesis Under Different Environmental Conditions, Springer, Progress in Photosynthesis Research, edited by: Biggins, J., 221–224, https://doi.org/10.1007/978-94-017-0519-6_48, 1987.

Bergh, J. and Linder, S.: Effects of soil warming during spring on photosynthetic recovery in boreal Norway spruce stands, Global Change Biol., 5, 245–253, https://doi.org/10.1046/j.1365-2486.1999.00205.x, 1999.

Bergh, J., Mcmurtrie, R., and Linder, S.: Climatic factors controlling the productivity of Norway spruce: A model-based analysis, Forest Ecol. Manag., 110, 127–139, https://doi.org/10.1016/S0378-1127(98)00280-1, 1998.

Bonan, G.: Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests, Science, 320, 1444–1449, https://doi.org/10.1126/science.1155121, 2008.

Bréda, N., Cochard, H., Dreyer, E., and Granier, A.: Water transfer in a mature oak stand (Quercuspetraea): seasonal evolution and effects of a severe drought, Can. J. Forest Res., 23, 1136–1143, https://doi.org/10.1139/x93-144, 1993.

Böttcher, K., Markkanen, T., Thum, T., Aalto, T., Aurela, M., Reick, C., Kolari, P., Arslan, A., and Pulliainen, J.: Evaluating Biosphere Model Estimates of the Start of the Vegetation Active Season in Boreal Forests by Satellite Observations, Remote Sens., 8, 1–31, https://doi.org/10.3390/rs8070580, 2016.

Chen, J., Govind, A., Sonnentag, O., Zhang, Y., Barr, A., and Amiro, B.: Leaf area index measurements at Fluxnet Canada forest sites, Agr. Forest Meteorol., 140, 257–268, https://doi.org/10.1016/j.agrformet.2006.08.005, 2006.

Cornuet, J.-M., Marin, J.-M., Mira, A., and Robert, C.: Adaptive Multiple Importance Sampling, Scand. J. Stat., 39, 798–812, https://doi.org/10.1111/j.1467-9469.2011.00756.x, 2012.

Duane, S., Kennedy, A., Pendleton, B., and Roweth, D.: Hybrid Monte Carlo, Phys. Lett. B, 195, 216–222, https://doi.org/10.1016/0370-2693(87)91197-X, 1987.

Egea, G., Verhoef, A., and Vidale, P.: Towards an improved and more flexible representation of water stress in coupled photosynthesis-stomatal conductance models, Agr. Forest Meteorol., 151, 1370–1384, https://doi.org/10.1016/j.agrformet.2011.05.019, 2011.

Farquhar, G., Caemmerer von, S., and Berry, J.: A Biochemical Model of Photosynthetic CO_2 Assimilation in Leaves of C3 species, Planta, 149, 78–90, https://doi.org/10.1007/BF00386231, 1980.

Franks, P. J., Bonan, G. B., Berry, J. A., Lombardozzi, D. L., Holbrook, N. M., Herold, N., and Oleson, K. W.: Comparing optimal and empirical stomatal conductance models for application in Earth system models, Global Change Biol., 24, 5709–5723, https://doi.org/10.1111/gcb.14445, 2018.

Friend, A. and Kiang, N.: Land surface model development for the GISS GCM: Effects of improved canopy physiology on simulated climate, J. Climate, 18, 2883–2902, https://doi.org/10.1175/JCLI3425.1, 2005.

Gao, Q., Zhao, P., Zeng, X., Cai, X., and Shen, W.: A model of stomatal conductance to quantify the relationship between leaf transpiration, microclimate and soil water stress, Plant Cell Environ., 25, 1373–1381, https://doi.org/10.1046/j.1365-3040.2002.00926.x, 2002.

Gao, Y., Markkanen, T., Thum, T., Aurela, M., Lohila, A., Mammarella, I., Kämäräinen, M., Hagemann, S., and Aalto, T.: Assessing various drought indicators in representing summer drought in boreal forests in Finland, Hydrol. Earth Syst. Sci., 20, 175–191, https://doi.org/10.5194/hess-20-175-2016, 2016.

Gao, Y., Markkanen, T., Aurela, M., Mammarella, I., Thum, T., Tsuruta, A., Yang, H., and Aalto, T.: Response of water use efficiency to summer drought in a boreal Scots pine forest in Finland, Biogeosciences, 14, 4409–4422, https://doi.org/10.5194/bg-14-4409-2017, 2017.

Gelman, A. and Rubin, D.: Inference from Iterative Simulation Using Multiple Sequences, Statist. Sci., 7, 457–472, https://doi.org/10.1214/ss/1177011136, 1992.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., and Rubin, D.: Bayesian Data Analysis, Chapman and Hall/CRC, 3rd Edn., 2013.

Groenendijk, M., Dolman, A., van der Molen, M., Leuning, R., Arneth, A., Delpierre, N., Gash, J., Lindroth, A., Richardson, A.D. Verbeeck, H., and Wohlfahrt, G.: Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data, Agr. Forest Meteorol., 1–17, https://doi.org/10.1016/j.agrformet.2010.08.013, in press, 2010.

Hagemann, S. and Stacke, T.: Impact of the soil hydrology scheme on simulated soil moisture memory, Clim. Dynam., 44, 1731–1750, https://doi.org/10.1007/s00382-014-2221-6, 2015.

Ikawa, H., Nakai, T., Busey, R., Kim, Y., Kobayashi, H., Nagai, S., Ueyama, M., Saito, K., Nagano, H., Suzuki, R., and Hinzman, L.: Understory CO_2, sensible heat, and latent heat fluxes in a black spruce forest in interior Alaska, Agr. Forest Meteorol., 214–215, 80–90, https://doi.org/10.1016/j.agrformet.2015.08.247, 2015.

Iwema, J., Rosolem, R., Rahman, M., Blyth, E., and Wagener, T.: Land surface model performance using cosmic-ray and point-scale soil moisture measurements for calibration, Hydrol. Earth Syst. Sci., 21, 2843–2861, https://doi.org/10.5194/hess-21-2843-2017, 2017.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J., Heimann, M., Hooker-Stroud, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E., Köstler, C., Mathieu, P.-P., Pinty, B., Reick, C., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., Terwisscha van Scheltinga, A., Vossbeck, M., Widmann, H., and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, J. Geophys. Res.-Biogeosci., 118, 1414–1426, https://doi.org/10.1002/jgrg.20118, 2013.

Kattge, J., Knorr, W., Raddatz, T., and Wirth, C.: Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models, Global Change Biol., 15, 976–991, https://doi.org/10.1111/j.1365-2486.2008.01744.x, 2009.

Kelliher, F., Lloyd, J., Arneth, A., Byers, J., McSeveny, T., Milukova, I., Grigoriev, S., Panfyorov, M., Sogatchev, A., Varlargin, A., Ziegler, W., Bauer, G., and Schulze, E.-D.: Evaporation from a central Siberian pine forest, J. Hydrol., 205, 279–296, https://doi.org/10.1016/S0022-1694(98)00082-1, 1998.

Knauer, J., Werner, C., and Zaehle, A.: Evaluating stomatal models and their atmospheric drought response in a land surface scheme: A multibiome analysis, J. Geophys. Res.-Biogeosci., 120, 1894–1911, https://doi.org/10.1002/2015JG003114, 2015.

Knorr, W.: Satellite Remote Sensing and Modelling of the Global CO_2 Exchange of Land Vegetation: A Synthesis Study, Max-Planck-Institut für Meteorologie Examensarbeit, 49, 1894–1911, 1997.

Knorr, W.: Annual and interannual CO_2 exchanges of the terrestrial biosphere: process-based simulations and uncertainties, Global Ecol. Biogeogr., 9, 225–252, https://doi.org/10.1046/j.1365-2699.2000.00159.x, 2000.

Knorr, W., Kaminski, T., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Carbon cycle data assimilation with a generic phenology model, J. Geophys. Res-Biogeosci., 115, G04017, https://doi.org/10.1029/2009JG001119, 2010.

Kolari, P., Lappalainen, H., Hänninen, H., and Hari, P.: Relationship between temperature and the seasonal course of photosynthesis in Scots pine at northern timberline and in southern boreal zone, Tellus B, 59, 542–552, https://doi.org/10.1111/j.1600-0889.2007.00262.x, 2007.

Kolari, P., Kulmala, L., Pumpanen, J., Launiainen, S., Ilvesniemi, H., Hari, P., and Nikinmaa, E.: CO_2 exchange and component CO_2 fluxes of a boreal Scots pine forest, Boreal Environ. Res., 14, 761–783, 2009.

Kozlowski, T. and Pallardy, S.: Acclimation and adaptive responses of woody plants to environmental stresses, Bot. Rev., 68, 270–334, https://doi.org/10.1663/0006-8101(2002)068[0270:AAAROW]2.0.CO;2, 2002.

Kropp, H., Loranty, M., Alexander, H., Berner, L., Natali, S., and Spawn, S.: Environmental constraints on transpiration and stomatal conductance in a Siberian Arctic boreal forest, J. Geophys. Res.-Biogeosci., 122, 761–783, https://doi.org/10.1002/2016JG003709, 2017.

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., and Richardson, A. D.: Constraining a global ecosystem model with multi-site eddy-covariance data, Biogeosciences, 9, 3757–3776, https://doi.org/10.5194/bg-9-3757-2012, 2012.

Lagergren, F. and Lindroth, A.: Transpiration response to soil moisture in pine and spruce trees in Sweden, Agr. Forest Meteorol., 112, 67–85, https://doi.org/10.1016/S0168-1923(02)00060-6, 2002.

Launiainen, S., Katul, G., Kolari, P., Lindroth, A., Lohila, A., Aurela, M., Varlagin, A., Grelle, A., and Vesala, T.: Do the energy fluxes and surface conductance of boreal coniferous forests in Europe scale with leaf area?, Global Change Biol., 22, 4096–4113, https://doi.org/10.1111/gcb.13497, 2016.

Leuning, R.: A critical appraisal of a combined stomatal-photosynthesis model for C_3 plants, Plant Cell Environ., 18, 339–355, https://doi.org/10.1111/j.1365-3040.1995.tb00370.x, 1995.

Leuning, R.: Temperature dependence of two parameters in a photosynthesis model, Plant Cell Environ., 25, 1205–1210, https://doi.org/10.1046/j.1365-3040.2002.00898, 2002.

Lin, Y.-A., Medlyn, B., Duursma, R., Prentice, I., Wang, H., Baig, S., Eamus, D., de Dios, V., Mitchell, P., Ellsworth, D., de Beeck, M., Wallin, G., Uddling, J., Tarvainen, L., Linderson, M.-L., Cernusak, L., Nippert, J., Ocheltree, T., Tissue, D., Martin-StPaul, N., Rogers, A., Warren, J., De Angelis, P., Hikosaka, K., Han, Q., Onoda, Y., Gimeno, T., Barton, C., Bennie, J., Bonal, D., Bosc, A., Löw, M., Macinins-Ng, C., Rey, A., Rowland, L., Setterfield, S., Tausz-Posch, S., Zaragoza-Castells, J., Broadmeadow,

M., Drake, J., Freeman, M., Ghannoum, O., Hutley, L., Kelly, J., Kikuzawa, K., Kolari, P., Koyama, K., Limousin, J.-M., Meir, P., Lola da Costa, A., Mikkelsen, T., Salinas, N., and Sun, W.: Optimal stomatal behaviour around the world, Nat. Clim. Change, 5, 459–464, https://doi.org/10.1038/nclimate2550, 2015.

Linkosalo, T., Heikkinen, J., Pulkkinen, P., and Mäkipää, R.: Fluorescence measurements show stronger cold inhibition of photosynthetic light reactions in Scots pine compared to Norway spruce as well as during spring compared to autumn, Front. Plant Sci., 13, 1–8, https://doi.org/10.3389/fpls.2014.00264, 2014.

Louis, J.-F.: A parametric model of vertical eddy fluxes in the atmosphere, Bound.-Layer Meteor., 17, 187–202, 1979.

Mäkelä, J.: Modified half-hourly FLUXNET dataset for 10 Boreal forest sites (CA-Obs,CA-Ojp,CA-Qfo,FI-Hyy,FI-Ken,FI-Let,FI-Sod,RU-Fyo,RU-Zot,US-Prr) [Data set], Zenodo, https://doi.org/10.5281/zenodo.3240954, 2019.

Mäkelä, A., Hari, P., Berninger, F., Hänninen, H., and Nikinmaa, E.: Acclimation of photosynthetic capacity in Scots pine to the annual cycle of temperature, Tree Physiol., 24, 369–376, https://doi.org/10.1093/treephys/24.4.369, 2004.

Mäkelä, J., Susiluoto, J., Markkanen, T., Aurela, M., Järvinen, H., Mammarella, I., Hagemann, S., and Aalto, T.: Constraining ecosystem model with adaptive Metropolis algorithm using boreal forest site eddy covariance measurements, Nonlin. Processes Geophys., 23, 447–465, https://doi.org/10.5194/npg-23-447-2016, 2016.

Martino, L., Elvira, V., Luengo, D., and Corander, J.: An Adaptive Population Importance Sampler: Learning From Uncertainty., IEEE Trans. Signal Proc., 63, 4422–4437, https://doi.org/10.1109/TSP.2015.2440215, 2015.

Medlyn, B., Duursma, R., Eamus, D., Ellsworth, D., Prentice, I., Barton, C., Crous, K., De Angelis, P., Freeman, M., and Wingate, L.: Reconciling the optimal and empirical approaches to modelling stomatal conductance, Global Change Biol., 17, 2134–2144, https://doi.org/10.1111/j.1365-2486.2010.02375.x, 2011.

Medlyn, B., De Kauwe, M., and Duursma, R.: New developments in the effort to model ecosystems under water stress, New Phytol., 212, 5–7, https://doi.org/10.1111/nph.14082, 2016.

Muukkonen, P., Nevalainen, S., Lindgren, M., and Peltoniemi, M.: Spatial occurrence of drought-associated damages in Finnish boreal forests: results from forest condition monitoring and GIS analysis, Boreal Environ. Res., 20, 172–180, 2015.

Nemani, R., Keeling, C., Hashimoto, H., Jolly, W., Piper, S., Tucker, C., Myneni, R., and Running, S.: Climate-Driven Increases in Global Terrestrial Net Primary Production from 1982 to 1999, Science, 300, 1560–1563, https://doi.org/10.1126/science.1082750, 2003.

Nobel, P. (Ed.): Physicochemical and environmental plant physiology, Academic Press, https://doi.org/10.1016/B978-0-12-374143-1.X0001-4, 1999.

Owen, A. and Yi, Z.: Safe and Effective Importance Sampling, J. Am. Stat. Assoc., 95, 135–143, https://doi.org/10.2307/2669533, 2000.

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A., Maignan, F., Chevallier, F., Ciais, P., and Prunet, P.: A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle, Geosci. Model Dev., 9, 3321–3346, https://doi.org/10.5194/gmd-9-3321-2016, 2016.

Post, H., Vrugt, J.A. amd Fox, A., Vereecken, H., and Hendricks Franssen, H.-J.: Estimation of Community Land Model parameters for an improved assessment of net carbon fluxes at European sites, J. Geophys. Res.-Biogeosci., 122, 661–689, https://doi.org/10.1002/2015JG003297, 2017.

Powell, T., Galbraith, D., Christoffersen, B., Harper, A., Imbuzeiro, H., Rowland, L., Almeida, S., Brando, P., Lola da Costa, A., Costa, M., Naomi M. Levine, N., Malhi, Y., Saleska, S., Sotta, E., Williams, M., Meir, P., and Moorcroft, P.: Confronting model predictions of carbon fluxes with measurements of Amazon forests subjected to experimental drought, New Phytol., 200, 350–365, https://doi.org/10.1111/nph.12390, 2013.

Raddatz, T., Reick, C., Korr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K.-G., Wetzel, P., and Jungclau, J.: Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century?, Clim. Dynam., 29, 565–574, https://doi.org/10.1007/s00382-007-0247-8, 2007.

Rannik, Ü., Peltola, O., and Mammarella, I.: Random uncertainties of flux measurements by the eddy covariance technique, Atmos. Meas. Tech., 9, 5163–5181, https://doi.org/10.5194/amt-9-5163-2016, 2016.

Raoult, N. M., Jupp, T. E., Cox, P. M., and Luke, C. M.: Land-surface parameter optimisation using data assimilation techniques: the adJULES system V1.0, Geosci. Model Dev., 9, 2833–2852, https://doi.org/10.5194/gmd-9-2833-2016, 2016.

Reick, C., Raddatz, T., Brovkin, V., and Gayler, V.: Representation of natural and anthropogenic land cover change in MPI-ESM, J. Adv. Model. Earth Syst., 5, 1–24, https://doi.org/10.1002/jame.20022, 2013.

Richardson, A., Hollinger, D., Burba, G., Davis, K., Flanagan, L., Katul, G., Munger, J., Ricciutio, D., Stoy, P., Suyker, A., Verma, S., and Wofsy, S.: A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes, Agr. Forest Meteorol., 136, 1–18, https://doi.org/10.1016/j.agrformet.2006.01.007, 2006.

Richardson, A., Mahecha, M., Falge, E., Kattge, J., Moffat, A., Papale, D., Reichstein, M., Stauch, V., Braswell, B., Churkina, G., Kruijt, B., and Hollinger, D.: Statistical properties of random CO_2 flux measurement uncertainty inferred from model residuals, Agr. Forest Meteorol., 148, 38–50, https://doi.org/10.1016/j.agrformet.2007.09.001, 2008.

Richardson, A., Anderson, R., Arain, M., Barr, A., Bohrer, G., Chen, G., Chen, J., Ciais, P., Davis, K., Desai, A., Dietze, M., Dragoni, D., Garrity, S., Gough, C., Grant, R., Hollinger, D., Margolis, H., Mccaughey, H., Migliavacca, M., Monson, R., Munger, J. W., Poulter, B., Raczka, B., Ricciuto, D., Sahoo, A., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J., and Xue, Y.: Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis, Global Change Biol., 18, 566–584, https://doi.org/10.1111/j.1365-2486.2011.02562.x, 2012.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM5. PART I: Model description, Max Planck Institute for Meteorology Report, 349, 1–127, available at: http://www.mpimet.mpg.de/fileadmin/publikationen/Reports/max_scirep_349.pdf (last access: 16 September 2019), 2003.

Schulze, E., Kelliher, F., Korner, C., Lloyd, J., and Leuning, R.: Relationships among Maximum Stomatal Conductance, Ecosystem Surface Conductance, Carbon Assimilation Rate, and Plant Nitrogen Nutrition: A Global Ecology Scaling Exercise, Annu. Rev. Ecol. Syst., 25, 629–662, 1994.

Scott, D. W.: Multivariate Density Estimation and Visualization, available at: http://EconPapers.repec.org/RePEc:zbw:caseps:200416 (last access: 16 September 2019), 2004.

Sellers, P.: Canopy reflectance, photosynthesis and transpiration, Int. J. Remote Sens., 6, 1335–1372, https://doi.org/10.1080/01431168508948283, 1985.

Thum, T., Aalto, T., Laurila, T., Aurela, M., Kolari, P., and Hari, P.: Parametrization of two photosynthesis models at the canopy scale in northern boreal Scots pine forest, Tellus, 59B, 874–890, https://doi.org/10.1111/j.1600-0889.2007.00305.x, 2007.

Trudinger, C., Raupach, M., Rayner, P., Kattge, J., Liu, Q., Pak, B., Reichstein, M., Renzullo, L., Richardson, A., Roxburgh, S., Styles, J., Wang, Y., Briggs, P., Barrett, D., and Nikolova, S.: OptIC project: An intercomparison of optimization techniques for parameter estimation in terrestrial biogeochemical models, J. Geophys. Res-Biogeosci., 112, G02027, https://doi.org/10.1029/2006JG000367, 2007.

Ueyama, M., Tahara, N., Iwata, H., Euskirchen, E., Ikawa, H., Kobyashi, H., Nagano, H., Nakai, T., and Harazono, Y.: Optimization if a biochemical model with eddy covariance measurements in black spruce forests of Alaska for estimating CO_2 fertilization effects, Agr. Forest Meteorol., 222, 98–111, https://doi.org/10.1016/j.agrformet.2016.03.007, 2016.

Veach, E. and Guibas, L.: Optimally Combining Sampling Techniques for Monte Carlo Rendering, SIGGRAPH 1995 Proceedings, 419–428, https://doi.org/10.1145/218380.218498, 1995.

Wang, K.-Y.: Canopy CO_2 exchange of Scots pine and its seasonal variation after four-year exposure to elevated CO_2 and temperature, Agr. Forest Meteorol., 82, 1–27, https://doi.org/10.1016/0168-1923(96)02342-8, 1996.

Xu, Z., Shimizu, H., Yagasaki, Y., Ito, S., Zheng, Y., and Zhou, G.: Interactive Effects of Elevated CO_2, Drought, and Warming on Plants, J. Plant Growth Regul., 32, 692–707, https://doi.org/10.1007/s00344-013-9337-5, 2013.

Zhou, S., Duursma, R., Medlyn, B., Kelly, J., and Prentice, I.: How should we model plant responses to drought? An analysis of stomatal and non-stomatal responses to water stress, Agric. Forest Meteorol., 182–183, 204–214, https://doi.org/10.1016/j.agrformet.2013.05.009, 2013.

Paper III

Uncertainty sources in simulated ecosystem indicators of the 21st century climate change

Mäkelä, J., Minunno, F., Aalto, T., Mäkelä, A., Markkanen, T., and Peltoniemi, M.

Biogeosciences Discuss., in review,
doi:10.5194/bg-2019-395, 2019.

©Authors 2019.
Reprinted under Creative Commons Attribution 4.0 License.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

Biogeosciences Discussions

Uncertainty sources in simulated ecosystem indicators of the 21st century climate change

Jarmo Mäkelä[1], Francesco Minunno[2], Tuula Aalto[1], Annikki Mäkelä[2], Tiina Markkanen[1], and Mikko Peltoniemi[3]

[1]Finnish Meteorological Institute, P.O. Box 503, FI-00101 Helsinki, Finland
[2]Department of Forest Sciences, P.O. Box 27, FI-00014 University of Helsinki, Finland
[3]Natural Resources Institute Finland, P.O. Box 2, FI-00791 Helsinki, Finland

Correspondence: Jarmo Mäkelä (jarmo.makela@fmi.fi), Francesco Minunno (francesco.minunno@helsinki.fi)

Abstract. The forest ecosystems are already responding to increased CO_2 concentrations and changing environmental conditions. These ongoing developments affect how societies can utilise and benefit from the woodland areas in the future, be it e.g. climate change mitigation as carbon sinks, lumber for wood industry or preserved for nature tourism and recreational activities. We assess the effect and the relative magnitude of different uncertainty sources in ecosystem model simulations from the year 1980 to 2100 for two Finnish boreal forest sites. The models used in this study are the land ecosystem model JSBACH and the forest growth model PREBAS. The considered uncertainty sources for both models are model parameters, four prescribed climates and two RCP (Representative Concentration Pathway) scenarios. PREBAS simulations also include an additional RCP scenario and two forest management actions. We assess the effect of these sources at four different stages of the simulations on several ecosystem indicators of climate change, e.g. gross primary production (GPP), ecosystem respiration, soil moisture, recurrence of drought, length of the vegetation active period (VAP), length of the snow melting period and the stand volume. The climate model uncertainty remains roughly the same throughout the simulations and is overtaken by the RCP scenario impact halfway through the experiment. The management actions are the most dominant uncertainty factors for Hyytiälä and as important as RCP scenarios at the end of the simulations, but contribute only half as much for Sodankylä. The parameter uncertainty is the most elusive to estimate due to non-linear and adverse effects on the simulated ecosystem indicators.

1 Introduction

The global atmospheric greenhouse gas concentrations are rising and inducing changes in land ecosystem carbon balances, water cycles and their seasonality. The rate of the expected concentration rise depends on human actions and the corresponding emission pathways chosen. The pathways presented in IPCC AR5 report (IPCC, 2014) lead to a radiative forcing of 2.6 W/m^2 to 8.5 W/m^2 in the year 2100. In addition to climate pathways connected to human actions, the variability in the IPCC climate projections is due to model differences and to internal variability in the climate system. Climate sensitivity has proven to be extremely difficult to constrain (Knutti and Sedláček, 2012). The multi-model spread in e.g. temperature and precipitation has not been narrowing during the last few years despite substantial model development (Eyring et al., 2019). However, narrowing the uncertainties should not be the only aim and sign of progress in climate modelling. Models improve as more processes

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

are described in detail, which may also introduce new unknown uncertainties. Thus it is important to study what are the contributions of different factors to the total uncertainty of examined variables, and how does the uncertainty evolve in the future.

The climate models provide drivers for the land ecosystem models. The predictions by land ecosystem models are affected by the driver uncertainties and by uncertainties related to the land surface model itself. Usually, only variability between different models is examined (see e.g. Friend et al., 2014; Nishina et al., 2015), and the uncertainty related to model parameters is not taken into account (Reyer et al., 2016). The unaccounted model processes can lead to significant underestimation of the overall uncertainty (Trugman et al., 2018). Furthermore, the spread in the uncertainty of the model outcome depends on the variable and region investigated. High latitude ecosystems are predicted to experience significant changes due to climate warming (Schaphoff et al., 2015). The change in seasonality of the ecosystems is predicted to manifest itself via decrease in snow cover duration, earlier soil thaw and later soil freeze and longer growing season (Dye and Tucker, 2003; McDonald et al., 2004; Barichivich and Caesar, 2012). The longer growing season and warmer temperatures are predicted to increase both ecosystem carbon uptake and respiration (Piao et al., 2008), while harmful extremes connected to heat, soil drought and soil excess water are also predicted to become more severe (Ruosteenoja et al., 2017). The evolution of net ecosystem exchange (NEE), defined as the difference between net ecosystem primary production (NPP) and heterotrophic respiration (R_h), is rather uncertain in future due to opposing drivers and may follow a trend towards net emissions or net uptake.

Forest management in Finland is a strong modifier of ecosystem carbon budgets and usually an unaccounted source of uncertainty in future predictions. The harvesting intensity defines the impact to the ecosystem carbon exchange (Korkiakoski et al., 2018). According to Kalliokoski et al. (2018), the future forest productivity was predicted to increase towards the end of the century. The climate model ensemble predictions were the dominant source of uncertainty for forest productivity, but closer to the end of century the role of emission pathways became more important. Estimation of future development of ecosystem carbon budgets together with impact factors such as management, seasonality and water conditions adds information to the whole ecosystem functioning. Assessment of uncertainties related to carbon budgets and growing season length together with water and snow conditions is important in estimating the forests ability to provide ecosystem services related to e.g. carbon sequestration, wood harvesting, maintaining habitats and promoting nature tourism (Snell et al., 2018; Holmberg et al., 2019).

Here we estimate how biomass, carbon, growing season, water and snow -related ecosystem indicators of climate change and their uncertainties progress in the future. We engage two ecosystem models at southern and northern boreal forest sites – JSBACH is developed to study land surface processes with closely coupled carbon balances and hydrology, while PREBAS is aimed to study carbon budgets with implementation of forest management. Both models have been previously calibrated for boreal ecosystems (Mäkelä et al., 2019; Minunno et al., 2019). We estimate the contribution of model parameter uncertainty, climate model variability, RCP pathway and management actions to the total uncertainty of these indicators. We apply canonical correlation analysis (CCA) to cross-correlate the uncertainty sources with the chosen ecosystem indicators. Finally, we aim to combine the model estimates to determine which are the dominant sources of uncertainty in future ecosystem projections.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

2 Materials and methods

We will first briefly introduce the sites and their characteristics, followed by the RCP scenarios and climate models used in this study as well as the models used to run the simulations. Next we describe our ecosystem indicators of climate change and define the methods used to analyse the simulations.

2.1 Sites

The sites used in this study are called Hyytiälä (FI-Hyy; 61°51′N, 24°17′E, 180 m a.s.l.) and Sodankylä (FI-Sod; 67°22′N, 26°38′E, 179 m a.s.l.); they are respectively located in southern and northern Finland and represent the southern and northern boreal pine forests. These sites can be characterised as Boreal evergreen needleleaf forests, where the dominant species is the Scots pine (*Pinus sylvestris*).

The Hyytiälä site (Kolari et al., 2009) was planted in 1962, after burning and mechanical soil preparation. The soil type is Haplic Podzol on glacial till. The site has an understory of Norway spruce (*Picea abies*) and few deciduous trees. The maximum measured all-sided leaf area index (LAI) for the Scots pine is 6.5 m^2/m^2, the average measured annual precipitation is 709 mm and temperature 2.9 °C.

The Sodankylä site (Thum et al., 2007) has been naturally regenerated after forest fires and hosts trees ranging from approximately 50 to 100 years of age. The soil type is fluvial sandy Podzol. The ground vegetation consists of lichens, mosses and ericaceous shrubs. The maximum measured LAI for the Scots pine is 3.6 m^2/m^2, as determined from forest inventories, the annual precipitation is 527 mm and temperature -0.4 °C.

2.2 RCP scenarios and climate models

We selected model runs of the fifth phase of the Coupled Model Intercomparison Project (CMIP5; Meehl et al., 2009; Taylor et al., 2012) following three representative concentration pathway (RCPs), that reach radiative forcing levels of 2.6, 4.5 and 8.5 W/m^2 by the end of the century (Moss et al., 2010; van Vuuren et al., 2011). Throughout the historical period that ends in 2005 the land cover data and the greenhouse gas concentrations corresponding different RCPs follow common trajectories (Meinshausen et al., 2011).

Climate data for years 1980-2100 was obtained from five global climate models (GCMs; CanESM2, CNRM-CM5, GFDL-CM3, HadGEM2-ES and MIROC5). The climate variables were bias corrected and further down-scaled to a $0.2° \times 0.1°$ longitude-latitude grid, similarly to Lehtonen et al. (2016); Holmberg et al. (2019). The bias correction methods are described in Räisänen and Räty (2013); Räty et al. (2014). The harmonised FMI meteorological data by Aalto et al. (2013) was used as reference.

The sub-set of five climate models was selected because of their good performance in reproducing current climate in Northern Europe and because they provided complete data sets for running impact models (Lehtonen et al., 2016). The five chosen models represent well the variation from current climate conditions (1981-2010) to the end of the ongoing century (2070-2099). The winter-time (i.e. December, January and February) precipitation in Finland for the five models in RCP4.5, covers

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.


the range of variability depicted by 24 out of 28 CMIP5 models investigated by Ruosteenoja et al. (2016). In summer the precipitation change range is generally narrower than in winter and the selected models cover the range of roughly half of the 28 CMIP5 models. Winter temperature change shows intermediate values among the 28 models and the range captures the ranges of change shown by 11 models. In summer the five model selection represents the range of change depicted by the upper half of the 28 models analysed by Ruosteenoja et al. (2016). Furthermore, the five climate models represent host institutes from different countries and from three continents: Asia, Europe and North-America. CO_2 concentrations from the RCPs 2.6, 4.5 and 8.5 increased monotonously through the calendar years reaching respective global means of 421, 538 and 936 ppm by the end of the century. PREBAS was run with results from all five climate models and three RCP scenarios, whereas JSBACH simulations included only RCP4.5 and RCP8.5 due to missing bias corrected climate variables. Moreover and for the same reason, JSBACH was not run with the HadGEM2-ES climate model for RCP8.5.

2.3 The JSBACH model

The JSBACH ecosystem model (Kaminski et al., 2013) is the land-surface component of the Earth system model of the Max Planck Institute for Meteorology (MPI-ESM). In these simulations, the model setup and parameter distributions are derived from Mäkelä et al. (2019). JSBACH is used uncoupled from the atmosphere, applying five layers within a multilayer soil hydrological scheme (Hagemann and Stacke, 2015) and utilising the BETHY model for canopy/stomatal conductance control (Knorr, 2000). Additionally, the model effectively uses only one plant functional type (PFT), coniferous evergreen trees.

The JSBACH model uncertainty is represented by a set of 100 parameter vectors, defined and described in more detail in Appendix A. The parameter distributions were derived from the simulations described in Mäkelä et al. (2019), where the model is calibrated and validated with site level measurements from 10 different evergreen needleleaf forests throughout the boreal zone (including Hyytiälä and Sodankylä). In order to avoid confusion with the climate models, the model uncertainty will be henceforth referred to as parameter uncertainty.

The JSBACH model initial state was derived from the end state of several thousand year long regional simulations that equilibrate the soil carbon storages. In addition, the simulations included a simulation specific spin-up period of 20 years to ensure adequate site level LAI and soil water storages. The spin-up was achieved by running the model through the first 20 years of simulation data, saving the state of the model variables and using them as the initial state for the 120-year long simulations. This type of spin-up introduces a discontinuity between the initial state and the driving climate but differences in the examined climate indicators should be negligible.

2.4 The PREBAS model

PREBAS (Valentine and Mäkelä, 2005; Peltoniemi et al., 2015; Minunno et al., 2019) is a simplified forest carbon and water balance model, which also considers forest growth and management. It calculates photosynthesis (GPP) using a light-use-efficiency (LUE) approach and ambient CO_2 concentration (Peltoniemi et al., 2015; Minunno et al., 2016). Daily GPP is influenced by soil moisture, radiation, temperature, vapour pressure deficit and precipitation. The model also calculates evapotranspiration (ET) and updates the water balance daily. Mean tree growth is calculated from GPP and respiration at an annual

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.


time step, and growth is allocated to different tree organs under assumptions on tree structure (Valentine and Mäkelä, 2005). The model includes tree mortality due to crowding. The growth module annually updates the canopy leaf area index (LAI) for the GPP and ET estimation. In order to estimate soil carbon, the annual litter fall is calculated by the growth allocation module, and fed to Yasso07 soil carbon model (Liski et al., 2005; Tuomi et al., 2009). NEE is calculated annually.

In addition to weather data, PREBAS requires information about the initial state of the simulated forest, defined as soil fertility class, stand basal area, mean height and mean diameter, at an appropriate spatial resolution. This information was extracted from the multisource forest inventory data maps (Tomppo et al., 2014; Mäkisara et al., 2016). The forest resource maps have a 16 m resolution and report the forest data for the year 2015. The model was initialised with forest data extracted for an area of 8×8 km square centered at the eddy covariance towers of Hyytiälä and Sodankylä.

In this study, two management scenarios were used in PREBAS simulations. The business as usual (BAU) scenario follows present forest management recommendations in Finland (Rantala et al., 2011), where trees have to be at least 24–30 cm diameter at breast height (dbh; 130 cm) and of age from 60–100 years before harvesting. The delayed ecosystem logging (DEL) scenario aims for the near term carbon sink increase by increasing the minimum harvesting diameter to 36 cm dbh.

2.5 Ecosystem indicators of climate changes

We study the uncertainty sources related to key biophysical indicators and their future development. Thus we ran the JSBACH and PREBAS models with different combinations of climate, RCP and management (only for PREBAS) scenarios with each realisation of the model parameterisations, resulting in approximately 2 000 site specific simulations for JSBACH and 6 000 for PREBAS. These simulations produced daily variables that were used to calculate the ecosystem indicators of climate change, presented in Table 1. We have included details on how we calculated the derived variables (number of dry days, start and end days of growing season and snow melting period) in Appendix B.

2.6 Analysis of results

We analyse the results by producing means, standard deviations and correlations of the model variables. This analysis is based on the annual values or averages over certain months (e.g. summer soil water) – one value per year. We utilise the Mann-kendall test (Mann, 1945; Kendall, 1975) to verify the existence of trend lines and kernel density estimation (KDE) to visualise the distribution of values (this approach can be viewed as a smoothed histogram).

We also carried out canonical correlation analysis (CCA) to quantify the impact of the different factors on the ecosystem indicators. The factors in this analysis are parametric uncertainty (par), climate models (clim) and RCP scenarios (rcp) for JSBACH and additionally management scenarios (man) for PREBAS. The indicators were averaged and divided into four consecutive 30-year long periods for both models: 1980-2009 (reference), 2010-2039 (interim), 2040-2069 (mid-century) and 2070-2099 (future). This produced single indicator values for each period and simulation (single instance of each factor) that were calculated for both sites separately.

CCA is a multivariate extension of correlation analysis that allows identifying linear relationships between two sets of variables (Hotelling and Pabst, 1936). We summarise the CCA results with the use of the redundancy index (Rd) that expresses

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

Table 1. Ecosystem indicators derived from the recorded values of the JSBACH and PREBAS simulations, separated into groups for the canonical correlation analysis. The group names relate to biomass distribution, ecosystem carbon exchange, length of the growing season, water cycle and snow melting period.

Indicator	Abb.	Units	JSB	PRE	Group
basal area	BA	m^2 / ha		x	Biomass
stand volume	V	m^3 / ha		x	Biomass
harvested volume	Vharv	m^3 / ha		x	Biomass
volume of dead trees	Vmort	m^3 / ha		x	Biomass
tree biomass	Biom	kg(C)		x	Biomass
tree litterfall	Lit	kg(C)		x	Biomass
leaf area index	LAI	m^2 / m^2		x	Biomass
gross growth	Growth	kg(C) / year		x	Biomass
gross primary production	GPP	g(C) / m^2 day	x	x	Carbon
net primary production	NPP	g(C) / m^2 day	x	x	Carbon
net ecosystem exchange	NEE	g(C) / m^2 day	x	x	Carbon
respiration (at)	Raut	g(C) / m^2 day	x	x	Carbon
soil carbon	Csoil	kg(C)		x	Carbon
start of growing season	SOS	DOY	x	x	Growth
end of growing season	EOS	DOY	x	x	Growth
length of growing season	VAP	days	x	x	Growth
evapotranspiration	ET	mm / day	x	x	Water
annual soil water	aSW	mm		x	Water
summer soil water	sSW	mm	x	x	Water
number of dry days	Ddry	days	x		Water
albedo	alb		x		Snow
snow amount	snow	m	x		Snow
start of snow melt	melt	DOY	x		Snow
snow clear date	clear	DOY	x		Snow
length of snow melt	SM	days	x		Snow

155 the amount of variance of a set of variables explained by another set of variables (Stewart and Love, 1968; Weiss, 1972; van den Wollenberg, 1977). The details of the CCA and the redundancy index are given in appendix C.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

3 Results

Forest management was the most dominant factor of uncertainty for Hyytiälä (Fig. 1) throughout the simulation. There was a clear difference for Sodankylä, where management gains only half as much influence. Disregarding management, the climate models and RCP scenarios represent major sources of both JSBACH and PREBAS predictive uncertainty. The impact of climate models was dominant during the reference and interim periods and remained roughly constant over time. The importance of RCP scenarios increased towards the end of the simulations, catching up to management impact at Hyytiälä in mid-century and representing the most important factor during the last period. The parametric uncertainty was the least influential factor for both JSBACH and PREBAS, at both sites.

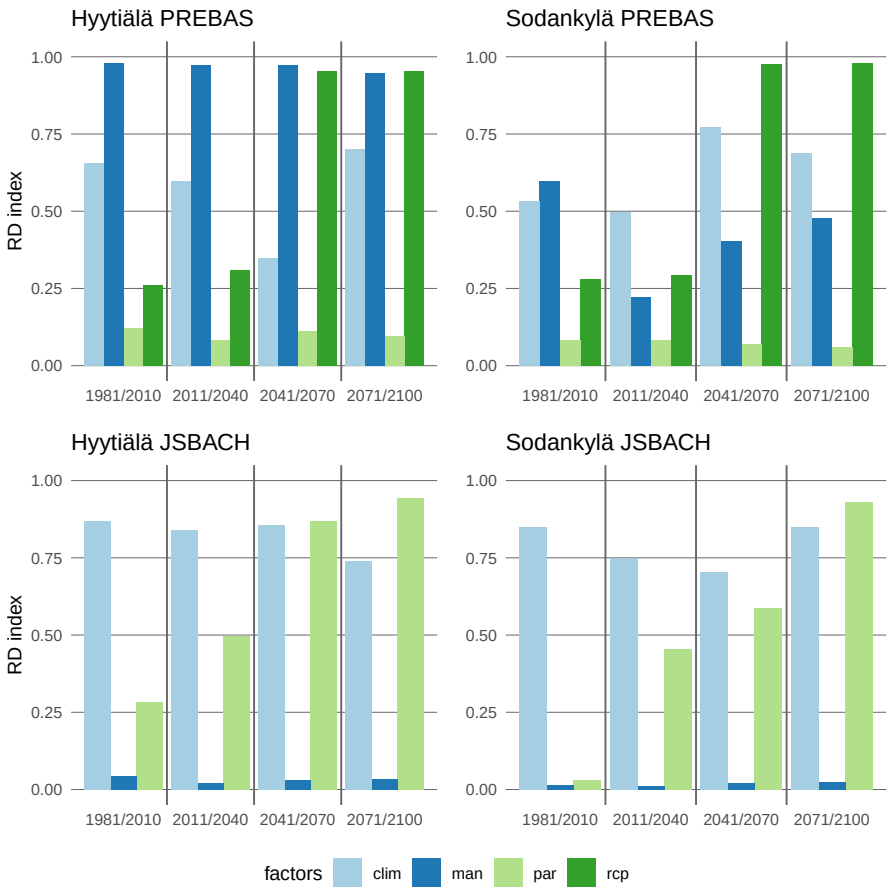

Figure 1. Redundancy indices calculated using all ecosystem indicators.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

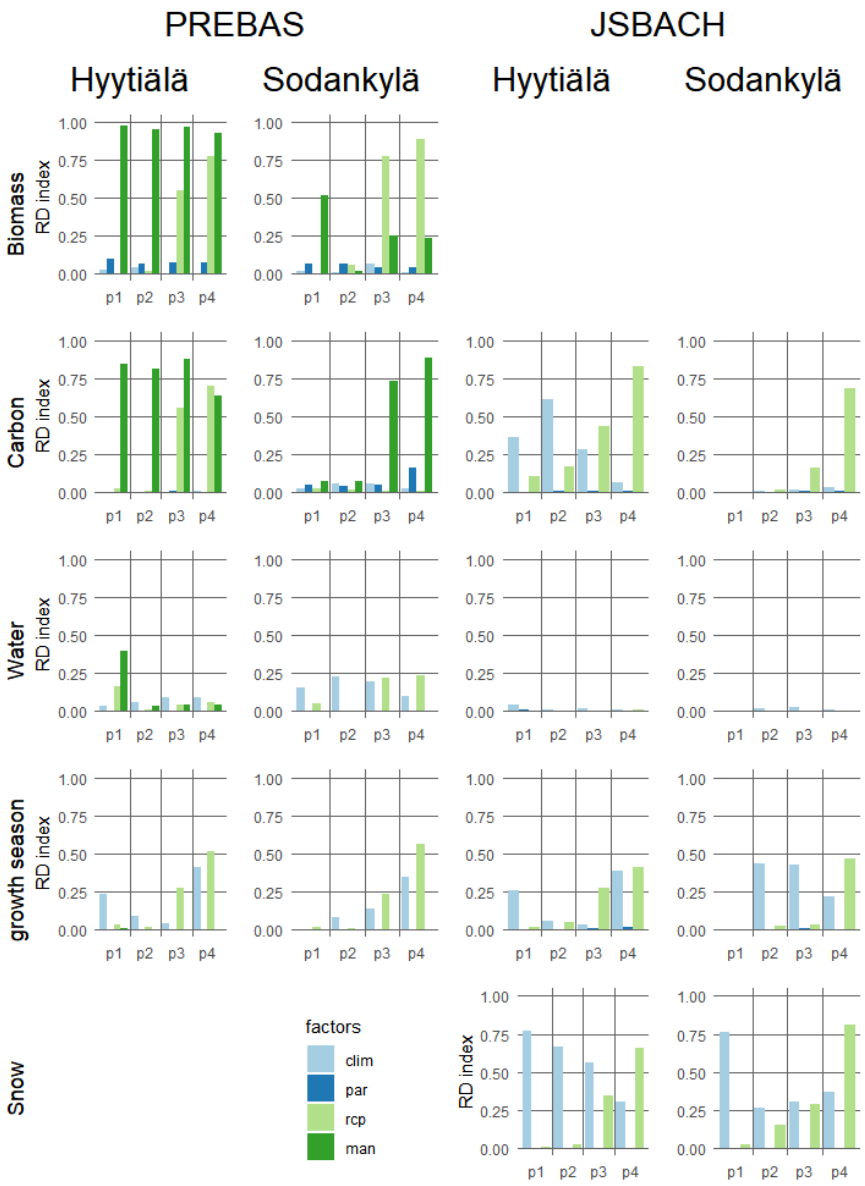

Figure 2. Redundancy indices calculated separately for the different indicator groups.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

3.1 Biomass distribution

The site-level differences in biomass stock uncertainties largely arise from the management actions (Fig. 2) and the management and RCP scenario impacts reflect the redundancy indices calculated with all ecosystem indicators (Fig. 1) for PREBAS. The RCP scenario influence increases for both sites towards the end of the simulations and the climate model and parameter uncertainty is negligible for both sites and all periods. There is an anomaly for Sodankylä reference period, where management has a very large impact. This situation arises due to minimal (0.1 m^3/ha), but systematic difference in harvested volume – the difference is so small it is not visually evident (Fig. 3). The rest of the Sodankylä reference period variables are nearly identical, so the small change in harvesting results in high correlation, which is captured by the CCA.

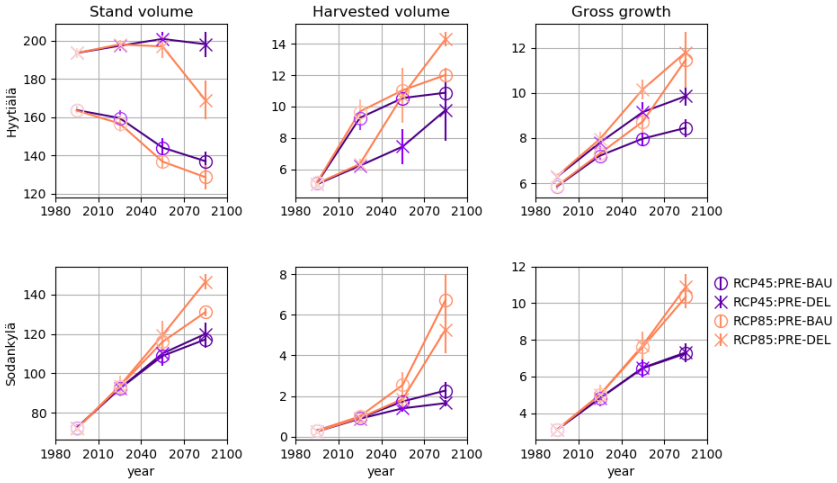

Figure 3. Selected ecosystem indicators from the PREBAS biomass factors, averaged for the 30-year long periods. The y-axis "whiskers" at each point represent the point specific uncertainty: one standard deviation amongst the corresponding simulations. We use lighter shading for the earlier periods, a different colour for the RCP scenarios and a different marker to separate the management actions.

The differences in site-specific variables due to the management actions, can already be seen from the reference period indicators (Fig. 3). The DEL scenario has approximately 10 % larger stand volume than BAU for Hyytiälä, but there is practically no difference for Sodankylä. The management actions start to have a noticeable impact for Sodankylä simulated variables at mid-century, but this impact is much smaller than that of the RCP scenarios. The management effect is much more pronounced at Hyytiälä, where both actions follow separate pathways.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

3.2 Ecosystem carbon exchange

The bifurcation of the annual GPP and respiration in JSBACH illustrates the separation of the RCP scenarios at about the midpoint (2040) in the simulations (Fig. 4). These two variables that comprise the net ecosystem exchange (NEE), have strong temporal linear correlations for both RCP scenarios ($r2 \approx 0.95$). The respective linear regression lines for GPP [g(C)/m^{2}d] yield an increase of 1.3 and 2.4 (RCP4.5 and 8.5) in 100 years for Hyytiälä and similarly 0.6 and 0.8 for Sodankylä. Likewise, the increases in respiration are 1.6 and 2.6 for Hyytiälä in 100 years and 0.8 and 1.2 for Sodankylä. GPP uncertainty was larger at the beginning of the simulations, but levelled with respiration at the end of the period. Relatively, the increased radiative forcing yields a stronger increase in GPP for Hyytiälä and respiration for Sodankylä. Some of the flux variables, such as Sodankylä GPP (Fig. 4), suggest a bi-modal value distribution in the the last 30 years of the simulations. This is caused by the different climate models yielding separate modes to the otherwise nearly identical value distributions. Most of the GPP and respiration value distribution (Fig. 4) reflect the variation in model parameterisations.

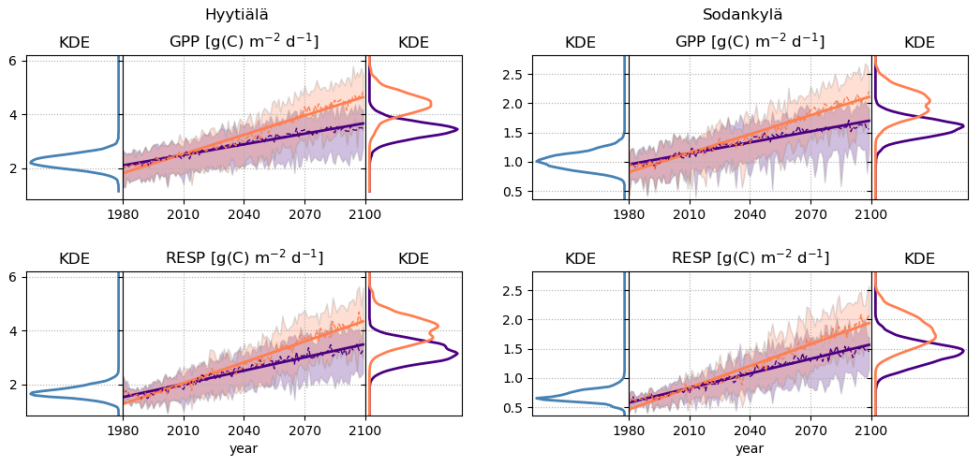

Figure 4. JSBACH predicted annual values of GPP and respiration for RCP4.5 (purple) and RCP8.5 (orange) scenarios. The shaded area represents all RCP-specific simulations, the dashed line is the annual mean and the solid line is the trend line. The KDE estimates on the left side of each image represents the distribution of the reference päeriod values of both RCP scenarios (blue), whereas the KDE on the right side consists of RCP specific values from the last 30 years of simulations.

As the bifurcating GPP and respiration fluxes signal, the RCP scenarios were important sources of uncertainty for the ecosystem carbon exchange variables at both sites, with importance growing over time (Fig. 2). However, it is noteworthy that management induced uncertainty for ecosystem carbon exchange was the most influential factor for Hyytiälä when it is accounted for in the model. The Sodankylä flux variation seems to be only dependent on the RCP scenario for both models, while the climate models were the most important factors at Hyytiälä during the first two periods for JSBACH.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

3.3 Ecosystem seasonality

The seasonal indicators depict the length of the vegetation active period and the snow melting period as well as the amount of soil water (and the recurrence of summer drought). The CCA analysis (Fig. 2) indicates that growing season indicators respond to changes in both climate models and RCP scenarios for both models, but the indicators are not sensitive to management actions. The snow melting period uncertainty for JSBACH is dominated by the climate models for the first half of the simulations for Hyytiälä, after which the RCP scenario is more influential. The situation is a bit different for Sodankylä snow melt, where the climate model uncertainty reduces radically after the reference period and then remains the same – the RCP scenarios gain effectiveness as simulations progress and reach the climate model influence at mid-century. The uncertainty related to the water balance for JSBACH is not explained by any of the examined factors and the uncertainties for PREBAS are also low.

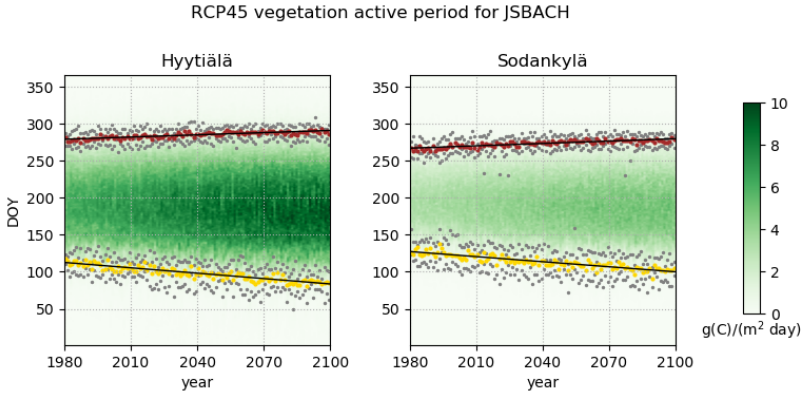

Figure 5. Average vegetation active period for JSBACH RCP4.5; yellow dots are the SOS values, red dots are the EOS values and the grey dots are the minimum and maximum SOS/EOS from all simulations. Also presented are the trend lines and the daily GPP as the green amplitude.

The vegetation active period is lengthening at both sites (Fig. 5). The displacement of the trendline start of (vegetation active) season (SOS) for JSBACH is approximately -8.1 days in 100 years for Hyytiälä (-11.3 for RCP8.5) and -7.6 days for Sodankylä (-10.9). Likewise, the end of season (EOS) displacement is 3.3 days for Hyytiälä (5.1 for RCP8.5) and 3.5 days for Sodankylä (5.2). The SOS and EOS temporal correlations are typically strong ($r^2 \approx 0.8$). The increase to the length of VAP is very similar for both sites, regardless of the different annual GPP.

The Mann-Kendall tests report a decreasing trend (earlier occurrence) for start of the snow melting period, first snow-free date and the length of the snow melting period (Fig. 6) in all simulations, except for Sodankylä RCP8.5 where the Mann-Kendall signifies the absence of trend for the melting period length. The simulations indicate that at the end of the century, the annual amount of snow in Hyytiälä will be radically diminished, and that Sodankylä winters will be similar to present

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

day Hyytiälä winters (especially in the RCP8.5 scenario). Relatively, the first snow free date is catching up to the start of the snow melting period (Fig. 6). The snow starts to melt approximately 20.7 days earlier in 100 years time for Hyytiälä RCP4.5 and 24.9 days earlier in RCP8.5, whereas the snow free dates appear 29.8 days (RCP4.5) and 41.7 days (RCP8.5) earlier. The corresponding values for Sodankylä are 12.2 (RCP4.5) and 25.1 (RCP8.5) for the start of snow melting period and 20.0 (RCP4.5) and 28.2 (RCP8.5) for the snow free dates. The correlations vary widely: $r^2 \approx 0.7$ for snow free dates, $r^2 \approx 0.5$ for the start of the melting period and $r^2 \approx 0.2$ for their difference.

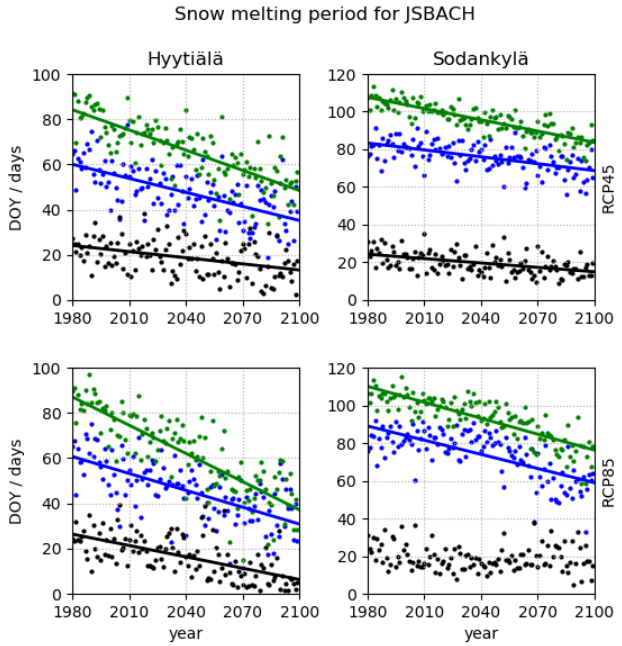

Figure 6. The average snow melting period for the JSBACH model; presented are the average annual values for the start of the snow melting period (blue), the first snow free day of the year (green) and their difference (black) as well as trend lines (calculated from the shown values) for these variables (when applicable).

The initial distributions of the summertime soil moisture values (Fig. 7) are unimodal for Hyytiälä and bimodal for Sodankylä for all climate models. This structure is still evident for the RCP4.5 scenario (of the last 30 years) but breaks down for the RCP8.5. Moreover, Hyytiälä RCP8.5 demonstrates some bimodality for two of the climate models whereas the RCP8.5 for Sodankylä seems to be losing the bimodality and is becoming (in appearance) more similar to the Hyytiälä reference period. The model parameterisations result in highly similar soil moisture distributions for the reference period, but there are clear differences (distribution modes and shapes) for the last 30 years.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

Biogeosciences
Discussions
Open Access
EGU

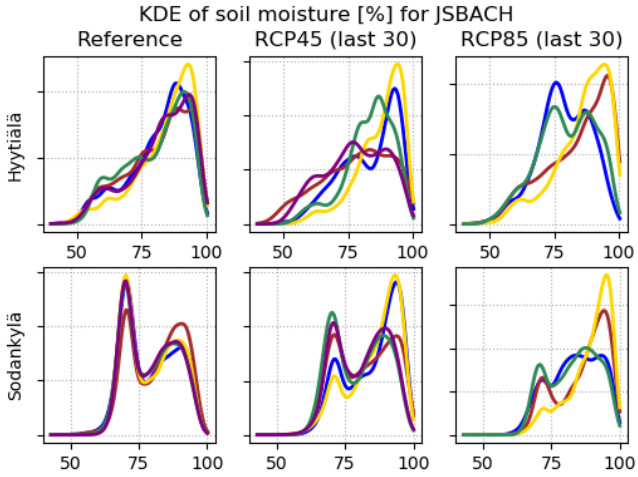

Figure 7. KDE estimates of the JSBACH soil moisture values (relative to soil field capacity) for the reference period and the last 30 years of simulations. Each colour represents the average summertime (June-August) soil moisture, produced with one of the climate models using all parameterisations.

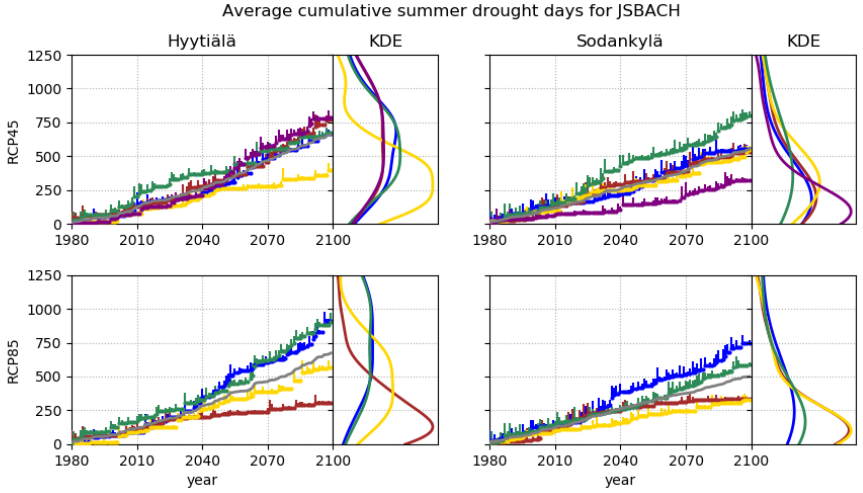

Figure 8. Accumulated summer drought days scatter plotted for each climate model, averaged over model parameterisations with minimum and maximum increment visualised as y-axis whiskers. The gray line is the average of the simulations. The KDE estimates on the right side depict the distribution of the accumulated drought days with the different parameterisations at the end of the simulation. The KDE figures have been cut at 1250 accumulated days.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.


The averaged drought events (Fig. 8) seem to be repeating at a roughly constant rate although the different model parameterisations result in wide soil moisture distributions (Fig. 7) at the end of the simulations. The average cumulative values correspond reasonably well with the drought indicator threshold in B1 (5 % of 92 summertime days, accumulated for 120 years would result in 552 days). The temporal correlations for the individual climate model and RCP specific simulations is poor (r^2 ranging from 0.12 to 0.5). The cumulative drought day distributions at the end of the simulations are strongly skewed with wide "tails" and high-value outliers (outside the figures) of approximately 2600 drought days for Hyytiälä and 3700 for Sodankylä. Interestingly, one of the climate models (brown) markedly reduces the amount drought days for the RCP8.5 at both sites when compared to RCP4.5. Neither the accumulated drought day variations or those of the soil moisture values (Fig. 7) are reflected in the CCA analysis of the Water group (Fig. 2)

3.4 Ecosystem indicator value comparison

The comparison results (Fig. 9) for soil moisture and ET indicate very small changes in the average values for both models but the JSBACH simulations manifest substantially larger variation. The JSBACH model yields more elevated levels of relative GPP, NPP, NEE and respiration for Hyytiälä, but the situation is (mostly) reversed for Sodankylä. These differences likely reflect the effect of the management actions and distinct site characteristics. The managements result in clearly different pathways for these variables at Hyytiälä, but only yield small differences at the end of the simulation for Sodankylä.

The SOS is roughly identical for both models, whereas both PREBAS versions have a larger effect on the EOS – initially the EOS for PREBAS occurs much earlier (roughly 15 days) than for JSBACH, which is diminished to a few days at the end of the simulations. The PREBAS extends the VAP more evenly from both "ends", whereas JSBACH focuses more on the SOS. These differences are reflected in the length of the VAP, which is merely the difference between EOS and SOS. Additionally, we note that the largest value spreads (deviations as represented by the length of the "whiskers") appear during the values representing the last 30 years of the RCP8.5 simulations – this merely reflects that the simulation uncertainties are increasing towards the end of the simulation (as expected). Overall, the model responses to the different inputs is very alike, which results in linear dependencies between the variables (Fig. 9).

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

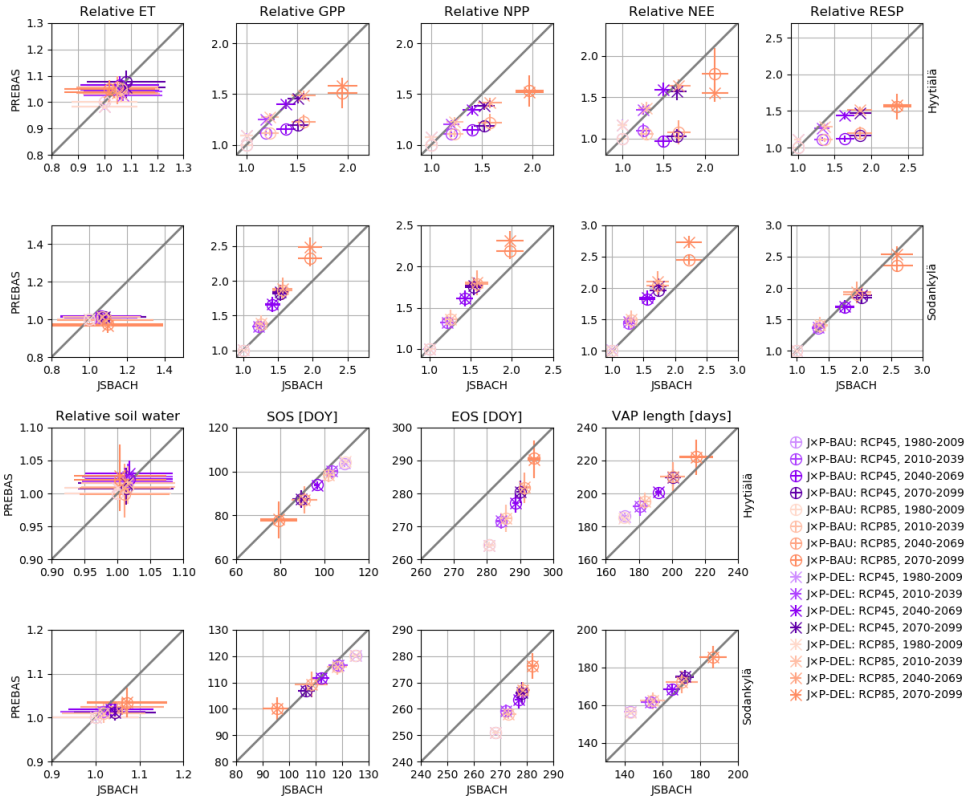

Figure 9. Average simulated values for shared ecosystem indicators between JSBACH and PREBAS, plotted for each 30-year period and both RCP4.5 and RCP8.5 scenarios. The values for JSBACH are divided by the average of the reference period values, and the values for PREBAS by the average of the BAU scenario reference period values. The "whiskers" at each point represent the point specific uncertainty: one standard deviation amongst the corresponding simulations. We use lighter shading for the earlier periods.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

Biogeosciences Discussions
Open Access
EGU

4 Discussion

In this paper we present an assessment on the importance of the different uncertainty sources, simulated on boreal forests for the 21st century. The JSBACH and PREBAS models yield similar uncertainty estimates (Fig. 1) and have a similar response to many of the examined ecosystem indicators of climate change (Fig. 9). The differences in modelled variables can be explained by the different model structures (e.g. soil moisture and evapotranspiration) or the inclusion of PREBAS management actions (ecosystem carbon fluxes). Forest management plays an important role in the estimates of ecosystem variables and their uncertainties. This importance is underscored by the lack of management in many land-surface components of climate models.

4.1 Impact to ecosystems

According to Grönholm et al. (2018), the long-term eddy-covariance measurements (1997–2017) at a boreal coniferous forest in Hyytiälä indicate a significant increase of gross-primary productivity (+10.5 [g(C)/m^2 year]), which is only partly compensated by an increased ecosystem respiration (+4.3g [g(C)/m^2 year]). As a result, the annual CO_2 sink has increased by about 6.2 [g(C)/m^2 year]. The GPP increase is dominated by an increase in LAI (from 4.1 to 4.6), while rise in the atmospheric CO2 concentration (from 360 ppm to 410 ppm) contributes only about 10 % to the rising GPP trend. It has to be noted that Hyytiälä forest was thinned in 2002, temporarily reducing LAI to 3.4. However, in few years the forest recovered to similar steadily increasing LAI trend than before thinning. The observed rise in the GPP is better replicated by the RCP8.5 scenario (Fig. 4) that yields an increase of +8.8 [g(C)/m^2 year] for Hyytiälä; whereas the increase in respiration is more closely reproduced by the RCP4.5 scenario (+5.8).

The RCP scenarios have a strong impact for growing stock and wood harvesting (Fig. 3), but the effect pales in comparison to the examined management actions. This underlines the importance of proper forest management for provisioning services (Snell et al., 2018; Holmberg et al., 2019). This is illustrated by the relative NEE pathways (Fig. 9) that are roughly convex for BAU and concave for DEL management actions. The simulations also indicate linearly lengthening VAP (Fig. 5), with high variation towards the end of the simulations (Fig. 9). This can be interpreted as beneficial for nature tourism and recreational activities, but on the other hand are the adverse effects of shortened snow melting period (Fig. 6) and potentially increased droughts (Fig. 8), also investigated by Ruosteenoja et al. (2017). These effects are also detrimental for winter harvesting and wood quality, as suggested by Holmberg et al. (2019).

Manninen et al. (2019) reported lengthened snow melting periods for some regions in Finland for 1982–2016. We analysed the reference period (1980–2009) snow melt in more detail and found that roughly 30 % of parameter specific simulations for Hyytiälä, and 20 % for Sodankylä, resulted in increased length for the snow melting period. We note that our simulations are restricted to site level, whereas regional experiments include lakes, rivers etc. that can significantly affect the outcome – this type of an uncertainty source is not considered in our simulations.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.


4.2 Simulation uncertainty sources

The overall uncertainty associated with the management actions differs for Hyytiälä and Sodankylä (Fig. 1). This is due to the more abundant harvesting effect at Hyytiälä (Fig. 3), whereas most of the biomass in Sodankylä is left to grow. Sodankylä stand volume increases as simulations progress whereas Hyytiälä stand volume remains the same or even decreases for the BAU scenario. This underlines the importance of proper forest management, as the impact of these relatively similar actions is strong – especially when taken into context of e.g. clear cuts.

As expected, the uncertainty related to the RCP scenarios increases systematically (Ruosteenoja et al., 2016) for all ecosystem indicators and grouped variables (except for the Water group) as the simulations advance further in time. This is similar to results by Kalliokoski et al. (2018). The RCP scenarios are the most dominant factor in explaining the JSBACH and PREBAS uncertainties for both sites at the end of the simulations. The RCP uncertainty also dominates the Carbon, Growth and Snow variables at both sites and Biomass variables for Sodankylä. The RCP scenarios tend to gain effect at mid-century (e.g. Fig. 3), although there are some earlier affects, e.g. snow variables for Sodankylä (Fig. 6).

The effect of the climate models to the redundancy indices is the most varied among the examined uncertainty sources. The climate models tend to have more impact in the two earlier periods, although the overall climate model uncertainty remains roughly the same throughout the simulations. This can be seen as arising from the internal variability of the climate system (Knutti and Sedláček, 2012). The combined variation of climate models and model parameters may not be fully captured due to non-linearity within the simulated variables. This is noted to emphasise the importance of the parameter uncertainty, as stated by Reyer et al. (2016). The parameter uncertainty is expected to be small when compared to the selected RCP scenarios that have a significant impact on the ecosystem (see Holmberg et al., 2019, Fig. 2). Most of the examined parameter distributions are highly alike for all climate models (Fig. 4), especially during the reference period (Fig. 7). The combined climate model and parameter uncertainty is on par with the RCP scenario uncertainty towards the end of the simulations (Fig. 1).

4.3 Validity of estimates

The JSBACH model calibration (Mäkelä et al., 2019) was originally used in comparison of various submodel components (stomatal conductance functions) and the PREBAS calibration (Minunno et al., 2019) utilised permanent growth and yield experiments. Both of these examinations rely on hindcasting with relatively recent meteorological measurements or datasets, and the resulting parameter distributions emulate the current climate conditions well. The examined ecosystem indicators were also chosen to reflect the calibrated parameters and processes.

The CCA analysis and model comparison focuses on the relative differences in the ecosystem indicators, and thus less importance is given to the absolute indicator values. The CCA analysis only accounts for linear dependencies (Hotelling and Pabst, 1936) between the input and output uncertainties, and even though the redundancy index (Stewart and Love, 1968) considers the (correlated) variance between the variables, the nonlinear effects may be underestimated. We reduce the annual variation and linearise the variables by averaging and separating them into four consecutive 30-year long periods. Additionally,

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.


we also examined the PREBAS redundancy indices without the RCP2.6 – these results differ only marginally from those with the RCP2.6 included, which increases the validity of the JSBACH results.

This linearisation may not be enough to capture all variation, as is the case with the JSBACH Water group uncertainties (Fig. 2) and the wide spread of soil moisture values (Fig. 7) and cumulative drought days (Fig. 8). The different parameterisations and climate models have a prominent variation, but due to adverse effects the correlations remain small. For example, the RCP8.5 radically increases precipitation (see Ruosteenoja et al., 2016, Fig. 2) and therefore increases the soil moisture (Fig. 7) and reduces the amount of drought days (Fig. 8). The strength of this effect varies among the climate models, but the model parameterisations still enable even radical increases to the number of drought days. This major source of uncertainty, investigated by e.g. Trugman et al. (2018), is not captured by CCA. However, when the indicators are reasonably correlated (as is the case for most of the presented indicators), the estimates are reliable.

The CCA analysis was performed for indicator groups to ensure robustness of the approach – this was not successful in every case, as a minimal but systematic difference in Sodankylä reference period harvested volume led to a large management scenario impact (Fig. 2). The situation arises as all of the other indicator values were nearly identical and thus a small systematic change that was relatively large, had high correlation and impact in CCA. This event was not replicated with the other groups.

5 Conclusions

Our simulations indicate that the management actions have the greatest influence to simulated ecosystem indicators of climate change. A similar impact is achieved by the RCP scenarios towards the end of century. The combined climate model and parameter uncertainty is also an important factor for the whole duration of the simulations due to internal variability of the climate system, but these effects can be easily underestimated due to non-linear or adverse effects. The examined uncertainties are comparable for both models.

Long-term measurements and simulations indicate considerable increases to GPP and respiration, with a slightly larger emphasis respectively for the southern and northern boreal forests. While the effect of management to these variables is linear, the impact on NEE is more complex and would be of interest in further studies. The snow melt is occurring several weeks earlier in all simulations and the length of the snow melting period appears to be decreasing, although the results for Sodankylä are not conclusive. Similarly, the length of the vegetation active period is expected to increase linearly for both sites by a few weeks. Sodankylä soil moisture is expected to increase, while the effects for Hyytiälä are varied. The scenarios do not constrain the recurrence of drought as the parameterisations enable varied outcomes.

Code and data availability. The JSBACH model can be obtained from the Max Planck Institute for Meteorology (MPI-M), where it is available for scientific community under the MPI-M Software License Agreement (http://www.mpimet.mpg.de/en/science/models/license/). The R package (Rprebas), containing the PREBAS model, is available on GitHub (https://github.com/checcomi/Rprebas). The model parameter values and the data used for the CCA analysis and redundancy index calculations are available as supplements to this paper.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

Appendix A: Model parameters

The JSBACH and PREBAS model uncertainties are represented by a set of parameter vectors (available as supplements). The different parameters and their distribution means and deviations are given in Tables A1 and A2. The PREBAS parameter values were evenly sampled from the MCMC chains in Minunno et al. (2019). The JSBACH parameter values were taken from the adaptive population importance sampler (APIS) simulations using the Bethy stomatal conductance formulation in Mäkelä et al. (2019). The bulk of these (100) vectors consists of APIS location parameters at 20 iterations (40 samples), which are complemented by later draws to reflect the sampling process.

Table A1. PREBAS model parameter descriptions as in Minunno et al. (Table 1; 2019) with distribution mean and standard deviation.

		pine		spruce		birch	
Parameter description (units)		μ	σ	μ	σ	μ	σ
Maintenance respiration rate of foliage (kg(C) kg^{-1}(C) yr^{-1}).	$m_{\mathrm{F,ref}}$	0.2	0.003	0.2	0.005	0.3	0.061
Maintenance respiration rate of fine roots (as above).	$m_{\mathrm{R,ref}}$	0.23	0.023	0.24	0.036	0.33	0.064
Maintenance respiration rate of sapwood (as above).	$m_{\mathrm{S,ref}}$	0.03	1.4×10^{-4}	0.03	3.0×10^{-4}	0.03	1.4×10^{-3}
Growth respiration rate (as above).	c	0.29	0.005	0.25	0.023	0.24	0.027
Leaf longevity (yr).	$\nu_{F,\mathrm{ref}}$	4.0	0.02	9.7	0.27	1.1	0.09
Fine root longevity (yr).	ν_{R}	0.9	0.03	1.7	0.07	1.2	0.19
Homogeneous extinction coefficient.	k_{H}	0.25	5.4×10^{-4}	0.25	8.8×10^{-4}	0.31	9.7×10^{-3}
Specific leaf area (m^2 kg^{-1}(C)).	s_{LA}	20.0	0.036	20.1	0.072	41.0	2.94
Parameter relating to reduction of photosynthesis with crown length.	s_1	0.011	6.1×10^{-4}	0.006	9.7×10^{-4}	0.031	0.011
Wood density (kg (C) m^{-3}).	ρ_{W}	197	2.82	183	2.48	226	20.9
Ratio of fine roots to foliage.	α_{Rs}	180	0.18	201	0.55	105	4.44
Foliage allometry parameter.	z	1.8	0.020	1.7	0.001	1.9	0.012
Ratio of total sapwood to above-ground sapwood biomass.	β_0	1.28	0.014	1.27	0.018	1.48	0.056
Ratio of mean branch pipe length to crown length.	β_{B}	0.4	4.5×10^{-4}	0.5	8.7×10^{-4}	0.4	0.048
Ratio of mean pipe length in stem above crown base to crown length.	β_{S}	0.39	0.006	0.46	0.007	0.46	0.024
Light level at crown base that prompts full crown rise.	C_{R}	0.22	0.008	0.16	0.004	0.17	0.013
Reineke parameter.	N_0	856	3.0	1040	7.4	998	68.6

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

Biogeosciences Discussions

Table A2. JSBACH model parameter descriptions as in Mäkelä et al. (Table 2; 2019) with distribution mean and standard deviation.

Parameter description (units)		μ	σ
Farquhar model maximum carboxylation rate at 25 °C (µmol (CO_2) m^{-2} s^{-1})	$V_{\mathrm{C,max}}$	42.8	1.94
Farquhar model efficiency for photon capture at 25 °C.	α	0.30	0.013
Multiplier in momentum and heat stability functions.	c_{b}	4.9	0.7
Ratio of unstressed C3-plant internal/external CO_2 concentration.	f_{C3}	0.81	0.025
Exponential scaling of water stress in reducing photosynthesis.	q	0.65	0.19
Volumetric soil water content above which fast drainage occurs.	θ_{dr}	0.79	0.09
Fraction depicting relative surface humidity based on soil dryness.	θ_{hum}	0.23	0.02
Volumetric soil moisture content at permanent wilting point.	θ_{pwp}	0.19	0.03
Volumetric soil moisture content, above which transpiration is unaffected.	θ_{tsp}	0.43	0.1
Fraction of precipitation intercepted by the canopy.	p_{int}	0.29	0.04
Depth for correction of surface temperature for snow melt (m).	s_{sm}	0.05	0.025
Maximum water content of the skin reservoir of bare soil (m).	w_{skin}	2.7×10^{-4}	7.3×10^{-5}
LoGro-P: memory loss parameter for chill days (days).	C_{decay}	15.7	5.3
LoGro-P: minimum value of critical heat sum (°C d).	S_{min}	18.0	6.4
LoGro-P: maximal range of critical heat sum (°C d).	S_{range}	189.0	49.9
LoGro-P: cutoff in alternating temperature (°C).	T_{alt}	6.0	1.8
LoGro-P: memory loss parameter for pseudo soil temperature (°C).	T_{ps}	15.8	5.3

Appendix B: Calculation of ecosystem indicators

Most of the ecosystem indicators in this paper are directly produced by the models, but few are derived from other variables.

B1 Drought days

The drought days are calculated as the amount of days when average soil moisture (of the combined 2nd and 3rd soil moisture levels in a 5-layer JSBACH scheme) is below a certain threshold. Only summertime (June, July, August) values are used and the threshold for Hyytiälä was set as the 5th percentile of all soil moisture values during the reference period. This value is approximately 33 % of the soil field capacity in Hyytiälä, which compares well with the parameters θ_{tsp} and θ_{pwp} for the Hyytiälä drought period optimisation in (Mäkelä et al., 2019). Thus, the number of dry days is a reasonable measure for

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.


Hyytiälä. We used the same percentile to set a similar value for Sodankylä although the site characteristics differ (different soil compositions and field capacity etc.).

B2 Vegetation active period

The dates for the start of season (SOS) and end of season (EOS) for the vegetative active period are calculated from simulated daily GPP. First we extracted the value corresponding to the 90th percentile of the daily GPP, from all of the simulations during the reference period, and then multiplied this value by 0.15. The SOS date is considered to be the first day of the year (DOY), when the daily GPP is consistently above this threshold. The consistency here means that, when we consider the daily GPP values, starting from the 30th DOY, to twice as far as the date of the SOS event, the GPP must be above the threshold for at least half of the days. The date for EOS is calculated similarly, when GPP is below the threshold and starting from 230th DOY.

B3 Snow melting period

The snow depth in model simulation varies on a year-to-year basis. We also encounter some years without any snow cover for Hyytiälä. Hence we first aggregate the snow depth over the model parameterisations and climate model simulations to produce average site and RCP scenario specific time series. This approach yields robust estimates of the snow cover, where the actual time series is smooth enough to allow calculation of the beginning of snow melting period and the first snow free date. We take a similar approach as in Manninen et al. (2019) and fit a sigmoidal function to identify the starting date of snow melt. The snow is considered to have melted, when the snow cover has consistently vanished. This means that there is no snow cover for at least half of the days during ± 10 days of the snow clear date.

Appendix C: Canonical correlation analysis

We carried out canonical correlation analysis (CCA) to quantify the impact of the different factors on the ecosystem indicators. These factors are parametric uncertainty (pars), management scenarios (man), climate models (clim) and rcp scenarios (rcp). CCA is a multivariate extension of correlation analysis that allows identifying linear relationships between two sets of variables (Hotelling and Pabst, 1936). It's use is similar to multiple regression but it is more appropriate when there are multiple intercorrelated variables such as model outputs. A more detailed description of CCA is provided in (Stewart and Love, 1968).

We consider two sets of variables, X (the different factors) and Y (ecosystem indicators). These are of dimensions N_p and N_q, where N is the number of realisations for the variable and p, q are the number of variables. In CCA we construct the linear composites (called canonical variates) $U_1 = a^T X$ and $V_1 = b^T Y$ by maximising the correlation between them. The composites U_1 and V_1 form the first pair of canonical variates. The second pair is formed similarly but it is required to be uncorrelated with the first pair (and so forth for the following pairs). The first pair accounts for the highest amount of variance between the two sets of variables and has the highest canonical correlation (Rc) – the variance and correlations diminish for each consecutive pair.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

The correlations between the individual variables (factors or indicators) and the respective canonical variates are called canonical loadings (CL), whereas the correlations with the opposite canonical variate are called canonical cross-loadings (CcL). These loadings are needed to summarise the CCA results via the use of the redundancy index (Rd) that expresses the amount of variance of a set of variables explained by another set of variables (Stewart and Love, 1968; Weiss, 1972; van den Wollenberg, 1977).

$$Rd_{iv} = \frac{1}{n_i}\left(\sum(CL_{iv}^2)\right) Rc_v^2 \tag{C1}$$

Above i is a placeholder for one of the two sets of variables, factors (f) and ecosystem indicators (e); v indicates a canonical variate; n_i is the number of variables in the i-th set and Rc are the canonical correlations.

The square of the canonical loadings expresses the proportion of variance accounted for each variable – computing the average for each variate provides an indication of the overall variability explained by the variate. The squared Rc represents the variance shared by the canonical variates of the two sets of variables – it is the bridge between the two sets. The redundancy index can be summed up across the canonical variates to have an overall measure of the bi-multivariate covariation of the two sets of variables.

In our analysis, we wanted to quantify the importance that each factor have on the ecosystem indicator uncertainty (RdF). We quantified the redundancy index of the indicators for each canonical variate and then multiplied it by the squared canonical cross-loadings between factors and variates.

$$RdF_{fv} = Rd_{ev} CcL_{fv}^2 \tag{C2}$$

CcL represents the proportion of variance shared between the factors (f) and the canonical variates of the ecosystem indicators (e). The RdF of the different factors can be summed up across the variates to obtain the overall weight that each factor has on the ecosystem indicator uncertainty.

Author contributions. J. Mäkelä and F. Minunno are respectively responsible for the JSBACH and PREBAS simulations. J. Mäkelä prepared the manuscript, with contributions from all co-authors, whereas F. Minunno provided the CCA and redundancy index calculations and analysis.

Competing interests. The authors declare that they have no conflicts of interest.

Acknowledgements. This work has been supported by Jenny and Antti Wihuri Foundation and the Ministry of foreign affairs of Finland (IBA-ForestFires, PC0TQ4BT-53) as well as the Academy of Finland Center of Excellence (272041), OPTICA (295874), the Strategic Research Council at the Academy of Finland (STN-SOMPA, 312932), ICOS Finland (281255) and MONIMET (LIFE12 ENV/FI/000409).

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

References

Aalto, J., Pirinen, P., Heikkinen, J., and Venäläinen, A.: Spatial interpolation of monthly climate data for Finland: comparing the performance of kriging and generalized additive models, Theor. Appl. Climatol., 112, 99–111, https://doi.org/10.1007/s00704-012-0716-9, 2013.

Barichivich, J., B. K. O. T. M. T. and Caesar, J.: Thermal growing season and timing of biospheric carbon uptake across the Northern Hemisphere, Glob. Biogeochem. Cycles, 26, GB4015, https://doi.org/10.1029/2012GB004312, 2012.

Dye, D. G. and Tucker, C. J.: Seasonality and trends of snow-cover, vegetation index, and temperature in northern Eurasia, Geophys. Res. Lett., 30, https://doi.org/10.1029/2002GL016384, 2003.

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., Collins, W. D., Gier, B. K., Hall, A. D., Hoffman, F. M., Hurtt, G. C., Jahn, A., Jones, C. D., Klein, S. A., Krasting, J. P., Kwiatkowski, L., Lorenz, R., Maloney, E., Meehl, G. A., Pendergrass, A. G., Pincus, R., Ruane, A. C., Russell, J. L., Sanderson, B. M., Santer, B. D., Sherwood, S. C., Simpson, I. R., Stouffer, R. J., and Williamson, M. S.: Taking climate model evaluation to the next level, Nat. Clim. Change, 9, 102–110, https://doi.org/10.1038/s41558-018-0355-y, 2019.

Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D. B., Dankers, R., Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R., Nishina, K., Ostberg, S., Pavlick, R., Peylin, P., Schaphoff, S., Vuichard, N., Warszawski, L., Wiltshire, A., and Woodward, F. I.: Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2, P. Natl. A. Sci. USA, 111, 3280–3285, https://doi.org/10.1073/pnas.1222477110, 2014.

Grönholm, T., Launiainen, S., Katul, G., Kolari, P., Aslan, T., Mammarella, I., Vesala, T., and Hari, P.: Does atmospheric CO2 explain increased carbon sink at a boreal coniferous forest flux site?, in: EGU General Assembly Conference Abstracts, vol. 20 of *EGU General Assembly Conference Abstracts*, p. 18561, 2018.

Hagemann, S. and Stacke, T.: Impact of the soil hydrology scheme on simulated soil moisture memory, Clim. Dynam., 44, 1731–1750, https://doi.org/10.1007/s00382-014-2221-6, 2015.

Holmberg, M., Aalto, T., Akujärvi, A., Arslan, A. N., Bergström, I., Böttcher, K., Lahtinen, I., Mäkelä, A., Markkanen, T., Minunno, F., Peltoniemi, M., Rankinen, K., Vihervaara, P., and Forsius, M.: Ecosystem Services Related to Carbon Cycling – Modeling Present and Future Impacts in Boreal Forests, Front. Plant Sci., 10, 343, https://doi.org/10.3389/fpls.2019.00343, 2019.

Hotelling, H. and Pabst, M. R.: Rank Correlation and Tests of Significance Involving No Assumption of Normality, The Annals of Mathematical Statistics, pp. 29–43, http://www.jstor.org/stable/2957508, 1936.

IPCC: Summary for Policymakers, Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, p. 151, [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)] IPCC, Geneva, Switzerland, https://www.ipcc.ch/report/ar5/syr/, 2014.

Kalliokoski, T., Mäkelä, A., Fronzek, S., Minunno, F., and Peltoniemi, M.: Decomposing sources of uncertainty in climate change projections of boreal forest primary production, Agr. Forest Meteorol., 262, 192–205, https://doi.org/10.1016/j.agrformet.2018.06.030, 2018.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J., Heimann, M., Hooker-Stroud, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E., Köstler, C., Mathieu, P.-P., Pinty, B., Reick, C., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., Terwisscha van Scheltinga, A., Vossbeck, M., Widmann, H., and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, J. Geophys. Res-Biogeo., 118, 1414–1426, https://doi.org/10.1002/jgrg.20118, 2013.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

Kendall, M.: Rank Correlation Methods, Charles Griffin, London, 4th edn., 1975.

Knorr, W.: Annual and interannual CO2 exchanges of the terrestrial biosphere: process-based simulations and uncertainties, Global Ecol. Biogeogr., 9, 225–252, https://doi.org/10.1046/j.1365-2699.2000.00159.x, 2000.

Knutti, R. and Sedláček, J.: Robustness and uncertainties in the new CMIP5 climate model projections, Nat. Clim. Change, 3, 369–373, https://doi.org/10.1038/nclimate1716, 2012.

Kolari, P., Kulmala, L., Pumpanen, J., Launiainen, S., Ilvesniemi, H., Hari, P., and Nikinmaa, E.: CO_2 exchange and component CO_2 fluxes of a boreal Scots pine forest, Boreal Environ. Res., 14, 761–783, http://www.borenv.net/BER/pdfs/ber14/ber14-761.pdf, 2009.

Korkiakoski, M., Tuovinen, J.-P., Penttilä, T., Sarkkola, S., Ojanen, P., Minkkinen, K., Rainne, J., Laurila, T., and Lohila, A.: Greenhouse gas and energy fluxes in a boreal peatland forest after clearcutting, Biogeosciences Discussions, 2018, 1–36, https://doi.org/10.5194/bg-2018-473, 2018.

Lehtonen, I., Kämäräinen, M., Gregow, H., Venäläinen, A., and Peltola, H.: Heavy snow loads in Finnish forests respond regionally asymmetrically to projected climate change, Natural Hazards and Earth System Sciences, 16, 2259–2271, https://doi.org/10.5194/nhess-16-2259-2016, 2016.

Liski, J., Palosuo, T., Peltoniemi, M., and Sievänen, R.: Carbon and decomposition model Yasso for forest soil, Ecol. Model., 189, 168–182, https://doi.org/10.1016/j.ecolmodel.2005.03.005, 2005.

Mäkelä, J., Knauer, J., Aurela, M., Black, A., Heimann, M., Kobayashi, H., Lohila, A., Mammarella, I., Margolis, H., Markkanen, T., Susiluoto, J., Thum, T., Viskari, T., Zaehle, S., and Aalto, T.: Parameter calibration and stomatal conductance formulation comparison for boreal forests with adaptive population importance sampler in the land surface model JSBACH, Geosci. Model Dev., 12, 4075–4098, https://doi.org/10.5194/gmd-12-4075-2019, 2019.

Mäkisara, K., Katila, M., Peräsaari, J., and Tomppo, E.: The multi-source national forest inventory of Finland – methods and results 2011, Nat. Resour. Bioeconomy Stud., pp. 1–215, http://urn.fi/URN:ISBN:978-952-326-186-0, 2016.

Mann, H. B.: Nonparametric Tests Against Trend, Econometrica, 13, 245–259, https://doi.org/10.2307/1907187, 1945.

Manninen, T., Aalto, T., Markkanen, T., Peltoniemi, M., Böttcher, K., Metsämäki, S., Anttila, K., Pirinen, P., Leppänen, A., and Arslan, A. N.: Monitoring changes in forestry and seasonal snow using surface albedo during 1982–2016 as an indicator, Biogeosciences, 16, 223–240, https://doi.org/10.5194/bg-16-223-2019, 2019.

McDonald, K., Kimball, J., Njoku, E., Zimmermann, R., and Zhao, M.: Variability in Springtime Thaw in the Terrestrial High Latitudes: Monitoring a Major Control on the Biospheric Assimilation of Atmospheric CO2 with Spaceborne Microwave Remote Sensing, Earth Interact., 8, 1–23, https://doi.org/10.1175/1087-3562(2004)8<1:VISTIT>2.0.CO;2, 2004.

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M. A., Greene, A. M., Hawkins, E., Hegerl, G., Karoly, D., Keenlyside, N., Kimoto, M., Kirtman, B., Navarra, A., Pulwarty, R., Smith, D., Stammer, D., and Stockdale, T.: Decadal Prediction, B. Am. Meteorol. Soc., 90, 1467–1486, https://doi.org/10.1175/2009BAMS2778.1, 2009.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic Change, 109, 213, https://doi.org/10.1007/s10584-011-0156-z, 2011.

Minunno, F., Peltoniemi, M., Härkönen, S., Kalliokoski, T., Makinen, H., and Mäkelä, A.: Calibration and validation of a semi-empirical flux ecosystem model for coniferous forests in the Boreal region, Ecol. Model., 341, 37–52, https://doi.org/10.1016/j.ecolmodel.2016.09.020, 2016.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

Minunno, F., Peltoniemi, M., Härkönen, S., Kalliokoski, T., Makinen, H., and Mäkelä, A.: Bayesian calibration of a carbon balance model PREBAS using data from permanent growth experiments and national forest inventory, Forest Ecol. Manag., 440, 208–257, https://doi.org/10.1016/j.foreco.2019.02.041, 2019.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., and Wilbanks, J. P. W. . T. J.: The next generation of scenarios for climate change research and assessment, Nature, 463, 747–756, https://doi.org/10.1038/nature08823, 2010.

Nishina, K., Ito, A., Falloon, P., Friend, A. D., Beerling, D. J., Ciais, P., Clark, D. B., Kahana, R., Kato, E., Lucht, W., Lomas, M., Pavlick, R., Schaphoff, S., Warszawaski, L., and Yokohata, T.: Decomposing uncertainties in the future terrestrial carbon budget associated with emission scenarios, climate projections, and ecosystem simulations using the ISI-MIP results, Earth System Dynam., 6, 435–445, https://doi.org/10.5194/esd-6-435-2015, 2015.

Peltoniemi, M., Pulkkinen, M., Aurela, M., Pumpanen, J., Kolari, P., and Mäkelä, A.: A semi-empirical model of boreal-forest gross primary production, evapotranspiration, and soil water – calibration and sensitivity analysis, Boreal Environ. Res., 20, 151–171, http://www.borenv.net/BER/pdfs/ber20/ber20-151.pdf, 2015.

Piao, S., Ciais, P., Friedlingstein, P., and Vesala, T.: Net carbon dioxide losses of northern ecosystems in response to autumn warming, Nature, 451, 49–52, https://doi.org/10.1038/nature06444, 2008.

Räisänen, J. and Räty, O.: Projections of daily mean temperature variability in the future: cross-validation tests with ENSEMBLES regional climate simulations, Clim. Dyn., 41, 1553–1568, https://doi.org/10.1007/s00382-012-1515-9, 2013.

Rantala, M., Leskinen, L., Hujala, T., and Kurttila, M.: Arvio METSO-ohjelman yhteistoimintaverkostohankkeiden vaikuttavuudesta ja kehittämistarpeista, Working Papers of the Finnish Forest Research Institute, 202, www.metla.fi/julkaisut/workingpapers/2011/mwp202.htm, 2011.

Räty, O., Räisänen, J., and Ylhäisi, J. S.: Evaluation of delta change and bias correction methods for future daily precipitation: intermodel cross-validation using ENSEMBLES simulations", journal="Climate Dynamics, Clim. Dyn., 42, 2287–2303, https://doi.org/10.1007/s00382-014-2130-8, 2014.

Reyer, C., Flechsig, M., Lasch-Born, P., and Van Oijen, M.: Integrating parameter uncertainty of a process-based model in assessments of climate change effects on forest productivity, Clim. Change, 137, 395–409, https://doi.org/10.1007/s10584-016-1694-1, 2016.

Ruosteenoja, K., Jylhä, K., and Kämäräinena, M.: Climate Projections for Finland Under the RCP Forcing Scenarios, Geophysica, 51, 17–50, 2016.

Ruosteenoja, K., Markkanen, T., Venäläinen, A., Räisänen, P., and Peltola, H.: Seasonal soil moisture and drought occurrence in Europe in CMIP5 projections for the 21st century, Clim. Dynam., 50, 1177–1192, https://doi.org/10.1007/s00382-017-3671-4, 2017.

Schaphoff, S., Reyer, C., Schepaschenko, D., Gerten, D., and Shvidenko, A.: Tamm Review: Observed and projected climate change impacts on Russia's forests and its carbon balance, Forest Ecol. Manag., 361, 432–444, https://doi.org/10.1016/j.foreco.2015.11.043, 2015.

Snell, R. S., Elkin, C., Kotlarski, S., and Bugmann, H.: Importance of climate uncertainty for projections of forest ecosystem services, Reg. Environ. Change, 18, 2145–2159, https://doi.org/10.1007/s10113-018-1337-3, 2018.

Stewart, D. and Love, W.: A general canonical correlation index, Psychol. Bull., 70, 160–163, https://doi.org/10.1037/h0026143, 1968.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, B. Am. Meteorol. Soc., 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

https://doi.org/10.5194/bg-2019-395
Preprint. Discussion started: 12 November 2019
© Author(s) 2019. CC BY 4.0 License.

Thum, T., Aalto, T., Laurila, T., Aurela, M., Kolari, P., and Hari, P.: Parametrization of two photosynthesis models at the canopy scale in northern boreal Scots pine forest, Tellus, 59B, 874–890, https://doi.org/10.1111/j.1600-0889.2007.00305.x, 2007.

Tomppo, E., Katila, M., Mäkisara, K., and Peräsaari, J.: The Multi-Source National Forest Inventory of Finland – Methods and Results 2011, Working Papers of the Finnish Forest Research Institute, pp. 1–224, http://urn.fi/URN:ISBN:978-951-40-2516-7, 2014.

Trugman, A. T., Medvigy, D., Mankin, J. S., and Anderegg, W. R. L.: Soil Moisture Stress as a Major Driver of Carbon Cycle Uncertainty, Geophys. Res. Lett., 45, 6945–6503, https://doi.org/10.1029/2018GL078131, 2018.

Tuomi, M., Thum, T., Järvinen, H., Fronzek, S., Berg, B., Harmon, M., Trofymow, J., Sevanto, S., and Liski, J.: Leaf litter decomposition – Estimates of global variability based on Yasso07 model, Ecol. Model., 220, 3362–3371, https://doi.org/10.1016/j.ecolmodel.2009.05.016, 2009.

Valentine, H. T. and Mäkelä, A.: Bridging process-based and empirical approaches to modeling tree growth, Tree Physiol., 25, 769–779, https://doi.org/10.1093/treephys/25.7.769, 2005.

van den Wollenberg, A.: Redundancy analysis an alternative for canonical correlation analysis, Psychometrika, 42, 207–219, https://doi.org/10.1007/BF02294050, 1977.

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., and Rose, S. K.: The representative concentration pathways: an overview, Climatic Change, 109, 5–31, https://doi.org/10.1007/s10584-011-0148-z, 2011.

Weiss, D. J.: Canonical correlation analysis in counseling psychology research, J. Couns. Psychol., 19, 241–252, https://doi.org/10.1037/h0032675, 1972.

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

FINNISH METEOROLOGICAL INSTITUTE
Erik Palménin aukio 1
P.O. Box 503
FI-00560 HELSINKI
tel. +358 29 539 1000
WWW.FMI.FI

FINNISH METEOROLOGICAL INSTITUTE
CONTRIBUTIONS No. 160
ISSN 0782-6117
ISBN 978-952-336-094-5 (paperback)
ISBN 978-952-336-095-2 (pdf)
https://doi.org/10.35614/isbn.9789523360952
Helsinki, 2020

Edita Prima Oy